

# Identification of key genes controlling docetaxel resistance in patients with breast cancer

Yannan Jiang (✉ [nicy1989@163.com](mailto:nicy1989@163.com))

Suzhou Municipal Hospital <https://orcid.org/0000-0002-9495-1813>

Yi Qian

Suzhou Municipal Hospital

Weifeng Qian

Suzhou Municipal Hospital

---

## Research article

**Keywords:** Breast cancer; Chemotherapy; Docetaxel; WGCNA

**Posted Date:** March 27th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-18265/v1>

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

[Read Full License](#)

---

# Abstract

Breast cancer is the most common malignant tumor in women. Chemotherapy is a very important part of its comprehensive treatment. Docetaxel is a commonly used chemotherapy drug in breast cancer treatment. However, some patients will develop drug resistance during use. To further explore this issue, we conducted an in-depth analysis on gene expression data of 24 breast cancer patients from GEO. Key module and hub genes were searched through WGCNA network construction, KEGG and GO function analysis, PPI network construction and other methods. Functional analysis revealed that the key module mainly related to drug transport events such as phagocytosis, cell-cell adhesion, and extracellular exosome. TCGA data were used to further verify these hub genes. Finally, we screened out 4 most important key genes, RPS2, EEF2, RPL15 and RPLP1, from 55 hub genes. These findings may highlight some therapeutic targets for predicting, avoiding or overcoming docetaxel resistance.

## Introduction

Breast cancer is the world's leading malignant tumor in women [1]. Although the early diagnosis, early treatment and individualized treatment have significantly reduced the mortality rate of breast cancer, its mortality rate still ranks among the highest female malignant tumors. Surgical resection is the radical treatment of breast cancer, but with the deepening of the understanding of breast cancer disease, the treatment of breast cancer patients has shifted from simple surgery to a holistic treatment combining surgery, chemotherapy, radiotherapy, targeted therapy and endocrine therapy[2–5]. Docetaxel is a kind of commonly used drugs in the chemotherapy for breast cancer, an M phase cell cycle specific drug, can promote tubular polymerization stability of microtubules, and suppress the depolymerization, leading to significantly reduction of small tube number and inhibition of cell proliferation. Moreover, docetaxel can destroy microtubules mesh structure, promote the cell apoptosis, thereby play an important role of anti-tumor in breast cancer[6]. Although docetaxel is effective in most breast cancer patients, some might eventually recur due to rapid proliferation of tumor cells and drug resistance, resulting in high mortality after disease progression[7–9]. Therefore, finding the key genes and potential pathways of docetaxel resistance in breast cancer is of great significance for individualized treatment and prognosis determination of breast cancer patients, and may also provide new therapeutic targets for future clinical practice.

The rapid development of gene microarray and sequencing technology has provided a new direction for the study of various cancers, especially the related pathological mechanism and the search for new molecular targets[10, 11]. At present, most studies focus on screening for differentially expressed genes without paying sufficient attention to the association between genes. The development of tumors is due to the interaction of multiple genes. Therefore, biological analysis methods that can cover a large amount of genomic data have been widely accepted and applied. Weighted correlation network analysis (WGCNA) establishes the connection between gene modules and clinical characters through scientific algorithms, establishes the regulatory network between gene concentration genes, and finally determines the key regulatory genes[12]. It is a mature analytical technique in gene data analysis and has been used

to find hub gene in a variety of cancers. For example, Zhou et al. identified 10 hub genes associated with the development and prognosis of pancreatic cancer using WGCNA[13]. Wu et al. also employed WGCNA to discover the potential biomarkers to differentiate malignant thyroid nodules[14].

In this study, we used gene expression data of 24 breast cancer patients from GEO database to construct WGCNA network and search for hub genes related to docetaxel resistance. We also performed functional analysis of the key module. In summary, we used a novel approach to identify the genes and possible pathways associated with docetaxel resistance in breast cancer.

## Methods

- Data collection

Gene chip data GSE349 and GSE350[7, 15] were downloaded from the public GEO database ([www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)), including 10 patients sensitive to docetaxel treatment, exhibiting less than 25% residual tumor, and 14 patients resistant to docetaxel treatment, exhibiting residual tumor of 25% or greater remaining volume. All samples were tested on the platform of Affymetrix Human Genome U95 Version 2 Array.

- WGCNA network construction

We calculated the differentially expressed genes in the docetaxel resistant group, compared with the sensitive group, and ranked them in descending order of significance. Weighted gene correlation network analysis (WGCNA) was performed using the WGCNA R package with the first 5,000 genes. Soft thresholding power  $\beta$  defined the scale-free topology fit index and the mean connectivity. We selected a  $\beta$  value to build the co-expression similarity and adjacency, then turned adjacency into topological overlap matrix (TOM), and made the hierarchical clustering tree on TOM-based dissimilarity. The minimum module size of the gene group was 30. Then we applied the module eigengenes (ME) function to get the correlations of different modules and sample traits. The highest one is the key module. At last, the module membership (MM)  $>0.8$  and the gene significance (GS)  $>0.8$  were defined as the thresholds for the selection of hub genes.

- Functional enrichment analysis

The DAVID website was selected to perform KEGG analysis and GO functional annotation to investigate and visualize the biological function of the key module (<https://david.ncifcrf.gov/tools.jsp>). The ggplot package in R was used for plotting. A p-value  $\leq 0.05$  and an FDR  $\leq 0.05$  were considered statistically significant.

- Protein-protein interaction (PPI) network analysis

The STRING website was selected to perform PPI network analysis to explore the relationships of the hub genes (<https://string-db.org/>). Nodes with higher PPI scores were considered to be more important key

genes.

- Overall Survival analysis

The Kaplan-Meier Plotter website was selected to perform overall survival analysis

(<https://kmplot.com/analysis/>). TCGA database is the source of information for this website. A p-value  $\leq$  0.05 was considered statistically significant.

## Results

- Construction of co-expression network

24 breast cancer samples used in this study to construct the WGCNA network all contained partial clinical information (Supplementary Table 1). According to the variation of gene expression, the 5,000 genes with the most significant changes were included in the follow-up study. First, we determined the optimal soft threshold ( $\beta$ ). In this study, the scale-free network construction standard is set as:  $\beta = 9$  (Figure 1). Based on this, network construction and gene module division were carried out. A total of 15 gene modules were distinguished by dynamic shear tree method, with different colors representing different gene modules and gray indicating that they did not belong to any gene module (Figure 1). Clinical information for breast cancer samples includes menopause, lymphatic metastasis, ER, PR, Her-2 status, and docetaxel resistance (demarcated by 25%). Pearson correlation coefficient (cor) was used to represent the correlation between module genes and clinical features, and the p-value was used to determine the significant correlation ( $P < 0.05$  represents significant correlation). Molecule Turquoise, Brown, Midnightblue, Red, Cyan, Greenyellow, Tan were correlated with drug resistance. Among them, the brown module was most correlated with docetaxel resistance (cor = 0.85,  $p < 1e-200$ ).

*Figure 1.* Construct of co-expression network and gene modules. A. The optimal soft threshold,  $\beta = 9$ . B. Distinguish 15 gene modules. C. Clinical information of breast cancer samples. D. Correlation between gene modules and clinical characters. E. Brown modules was most strongly associated with drug resistance (cor = 0.85,  $p < 1e-200$ ), including 735 genes.

- KEGG and GO functional analysis, and PPI network construction

To further understand the biological functions involved in brown module genes, KEGG and GO analyses were performed using DAVID online tool (Figure 2). The top 20 terms in GO analysis and the pathways ( $P < 0.05$ ) in KEGG analysis were listed in the Figure2 and Supplementary Table2. According to KEGG analysis, pathways associated with breast cancer or drug resistance included Phagocytosis, Estrogen signaling pathway and Pathways in cancer. The most prominent biological process in GO analysis is cell-cell adhesion. Cellular component is extracellular exosome. And molecular function is protein binding. Subsequently, the correlation coefficients of MM and GS were set as  $> 0.8$  (MM was the correlation between Gene and Module Brown, GS was the correlation between Gene and phenotype), and a total of 55 hub genes were found. Then PPI network was constructed (Figure 2E). The top 10 genes ranked by

score were RPS5, FBL, RPS2, EEF2, RPL29, RPLP1, RPS19, RPL4, RPL15 and EIF4G1. All these 10 genes were less expressed in docetaxel resistant breast cancer patients (Figure 3).

*Figure 2.* KEGG-GO-PPI analysis. A. Top 20 in GO (Biological Process). B. Top 20 in GO (Cellular Component). C. Top 20 in GO (Molecular Function). D. Pathways in KEGG analysis (p value $\leq$ 0.05). E. PPI network analysis with 55 hub genes.

- Overall survival analysis of hub gene

Breast cancer data from the TCGA database were used to verify the correlation between the top 10 hub genes and breast cancer patient survival. According to the results of gene chip comparison, FBL, RPS2, EEF2, RPLP1, RPS19, RPL4, and RPL15 were significantly correlated with overall survival (OS) (Supplementary Table 3). Combined with the expression of these genes in patient samples, RPS2, EEF2, RPL15 and RPLP1 seemed to be the most important key genes for docetaxel resistance in BC patients, and should be paid more attention to in future studies (Figure 3).

*Figure 3.* 10 Hub gene expression levels and their relationships to OS in BC patients. RPS2, EEF2, RPL15 and RPLP1 were significantly correlated with OS ( $P < 0.05$ ). The higher expression level indicated the better prognosis.

## Discussion

Breast cancer is the most common malignant tumor in women, accounting for about 24.2% of malignant tumors in women [1]. At present, the treatment of breast cancer patients has been a comprehensive treatment integrating surgery, chemotherapy, radiotherapy, targeted therapy, endocrine therapy and other treatment methods[2, 4, 16]. According to the immunohistochemical characteristics, it was clinically divided into Luminal A subtype, Luminal B subtype, HER-2 overexpression subtype and Basal-like subtype/triple negative breast cancer[17]. For patients who need chemotherapy, different chemotherapy regimens can be given to breast cancer patients according to different molecular types. Docetaxel is one of the most commonly recommended drugs. Docetaxel, a drug targeting microtubules, can affect the composition of cytoskeleton and the formation of mitotic spindles, so as to play an anti-tumor role in universities[15]. However, many patients eventually develop resistance to it. We need to conduct in-depth research into the reasons for this, discover the genes and underlying mechanisms that play an important role in drug resistance, explain the occurrence of this phenomenon, and screen and predict such patients. Moreover, these important genes are likely to become new clinical therapeutic targets.

In this study, we used chip data from 24 breast cancer patients from GEO database to construct WGCNA network. A total of 15 gene modules were found, among which the most related to Docetaxel resistance was the brown module, which included 735 genes. Through functional analysis of these genes, we found that the Pathways most associated with Docetaxel resistance were Phagocytosis, Estrogen signaling pathway and Pathways in cancer. These Pathways are mainly associated with oncogenesis and drug transport[18–21]. In addition, the cell-cell adhesion, extracellular exosome and protein binding were the

most prominent ones in the GO analysis. It will help us in our subsequent understanding of key genes and underlying mechanisms. Subsequently, 55 hub genes were screened out based on the correlation coefficients of MM and GS of  $>0.8$ . After PPI network construction, the most important 10 key genes were identified according to the scores, namely RPS5, FBL, RPS2, EEF2, RPL29, RPLP1, RPS19, RPL4, RPL15 and EIF4G1. Based on the survival information of breast cancer patients in TCGA database, we focused on 4 genes, RPS2, EEF2, RPL15 and RPLP1. The proteins encoded by these four genes are closely related to protein synthesis. Deregulation of ribosomal protein expression and translation has been reported to promote breast cancer growth and metastasis[22]. RPS2 (ribosomal 40S subunit protein S2), RPL15 (ribosomal protein L15) and RPLP1 (ribosomal protein lateral stalk subunit P1) are involved in ribosomal composition. EEF2 (eukaryotic translation elongation factor 2) encodes a member of the GTP-binding translation elongation factor family. RPS2 has been rarely reported in tumors, only in prostate cancer. RPS2 overexpression can promote the development of prostate cancer, partially through controlling let-7a expression [23, 24]. There are few studies about EEF2 function in tumor, but a lot about the effect of the EEF2 phosphorylation kinase, EEF2K. EEF2K can promote the growth and metastasis of triple negative breast cancer, as well as paclitaxel resistance of by regulating autophagy activity [25–27]. RPL15 is highly expressed in colorectal cancer tissues, affecting cell proliferation and apoptosis [28]. High expression of RPL15 in circulating tumor cells can also promote the metastasis of breast cancer [22]. RPLP1 was significantly increased in triple negative breast cancer tissues and associated with the tumor metastasis and patient prognosis [29]. According to the above reports, these four genes have not been directly studied in docetaxel resistance, but they all have a certain correlation with tumors. However, we found that this correlation is not completely consistent with the effect found in this study, which may be because of different tumor types, or because tumorigenesis and chemotherapy resistance are two different biological processes, and the reasons for this need to be further studied.

In summary, in this study, four key genes for docetaxel resistance in breast cancer were identified through GEO, TCGA public databases and WGCNA network construction. This has very important guiding significance to our clinical practice.

## Declarations

### Conflict of interest statement:

The authors declare that they have no competing interests.

### Acknowledgements:

This work was part of the Program on Basic Research for Application funded under Suzhou Health and Family Planning Commission, grant number SYS2019103.

The authors are grateful to all study participants.

### Author contribution statement:

Yannan Jiang designed experiments; Yannan Jiang and Yi Qian carried out experiments; Yannan Jiang analyzed experimental results. Weifeng Qian wrote the manuscript.

### Data availability Statement:

All data, models, or code generated or used during the study are available from the corresponding author by request.

## References

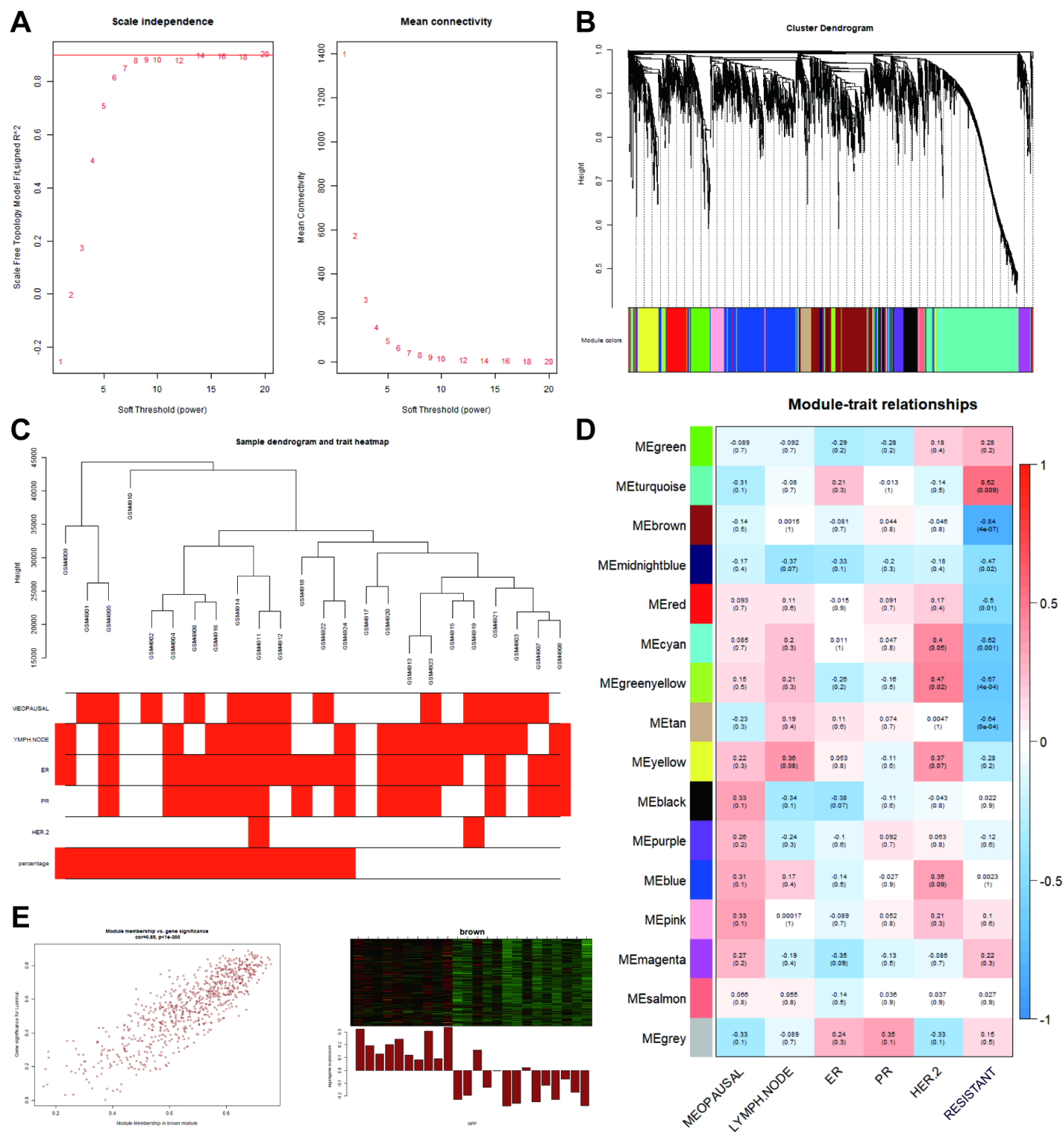
1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians* 2018, *68*(6):394–424.
2. Durrani S, Heena H: *Controversies Regarding Ovarian Suppression and Infertility in Early Stage Breast Cancer. Cancer management and research* 2020, *12*:813–817.
3. Spring LM, Wander SA, Andre F, Moy B, Turner NC, Bardia A: *Cyclin-dependent kinase 4 and 6 inhibitors for hormone receptor-positive breast cancer: past, present, and future. Lancet* 2020, *395*(10226):817–827.
4. Smith BD, Bellon JR, Blitzblau R, Freedman G, Haffty B, Hahn C, Halberg F, Hoffman K, Horst K, Moran J *et al: Radiation therapy for the whole breast: Executive summary of an American Society for Radiation Oncology (ASTRO) evidence-based guideline. Practical radiation oncology* 2018, *8*(3):145–152.
5. Mastectomy Decisions Audit Collaborative obotWMRC: *Multicentre prospective observational study evaluating recommendations for mastectomy by multidisciplinary teams. The British journal of surgery* 2020, *107*(3):227–237.
6. Jasra S, Anampa J: *Anthracycline Use for Early Stage Breast Cancer in the Modern Era: a Review. Current treatment options in oncology* 2018, *19*(6):30.
7. Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Tham YL, Kalidas M, Elledge R, Mohsin S, Osborne CK *et al: Patterns of resistance and incomplete response to docetaxel by gene expression profiling in breast cancer patients. Journal of clinical oncology: official journal of the American Society of Clinical Oncology* 2005, *23*(6):1169–1177.
8. Gomez-Miragaya J, Moran S, Calleja-Cervantes ME, Collado-Sole A, Pare L, Gomez A, Serra V, Dobrolecki LE, Lewis MT, Diaz-Lagares A *et al: The Altered Transcriptome and DNA Methylation Profiles of Docetaxel Resistance in Breast Cancer PDX Models. Molecular cancer research: MCR* 2019, *17*(10):2063–2076.
9. Gomez-Miragaya J, Diaz-Navarro A, Tonda R, Beltran S, Palomero L, Palafox M, Dobrolecki LE, Huang C, Vasaikar S, Zhang B *et al: Chromosome 12p Amplification in Triple-Negative/BRCA1-Mutated Breast*

- Cancer Associates with Emergence of Docetaxel Resistance and Carboplatin Sensitivity. Cancer research* 2019, 79(16):4258–4270.
- 10.Fakih M, Ouyang C, Wang C, Tu TY, Gozo MC, Cho M, Sy M, Longmate JA, Lee PP: *Immune overdrive signature in colorectal tumor subset predicts poor clinical outcome. The Journal of clinical investigation* 2019, 129(10):4464–4476.
- 11.Chen Y, Peng C, Chen J, Chen D, Yang B, He B, Hu W, Zhang Y, Liu H, Dai L *et al*: *WTAP facilitates progression of hepatocellular carcinoma via m6A-HuR-dependent epigenetic silencing of ETS1. Molecular cancer* 2019, 18(1):127.
- 12.Zhang B, Horvath S: *A general framework for weighted gene co-expression network analysis. Statistical applications in genetics and molecular biology* 2005, 4:Article17.
- 13.Zhou YY, Chen LP, Zhang Y, Hu SK, Dong ZJ, Wu M, Chen QX, Zhuang ZZ, Du XJ: *Integrated transcriptomic analysis reveals hub genes involved in diagnosis and prognosis of pancreatic cancer. Molecular medicine* 2019, 25(1):47.
- 14.Wu D, Hu S, Hou Y, He Y, Liu S: *Identification of potential novel biomarkers to differentiate malignant thyroid nodules with cytological indeterminate. BMC cancer* 2020, 20(1):199.
- 15.Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, Mohsin S, Osborne CK, Chamness GC, Allred DC *et al*: *Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. Lancet* 2003, 362(9381):362–369.
- 16.Shah AN, Metzger O, Bartlett CH, Liu Y, Huang X, Cristofanilli M: *Hormone Receptor-Positive/Human Epidermal Growth Receptor 2-Negative Metastatic Breast Cancer in Young Women: Emerging Data in the Era of Molecularly Targeted Agents. The oncologist* 2020.
- 17.Nielsen TO, Hsu FD, Jensen K, Cheang M, Karaca G, Hu Z, Hernandez-Boussard T, Livasy C, Cowan D, Dressler L *et al*: *Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clinical cancer research: an official journal of the American Association for Cancer Research* 2004, 10(16):5367–5374.
- 18.Saltarella I, Desantis V, Melaccio A, Solimando AG, Lamanuzzi A, Ria R, Storlazzi CT, Mariggio MA, Vacca A, Frassanito MA: *Mechanisms of Resistance to Anti-CD38 Daratumumab in Multiple Myeloma. Cells* 2020, 9(1).
- 19.Nigro A, Ricciardi L, Salvato I, Sabbatino F, Vitale M, Crescenzi MA, Montico B, Triggiani M, Pepe S, Stellato C *et al*: *Enhanced Expression of CD47 Is Associated With Off-Target Resistance to Tyrosine Kinase Inhibitor Gefitinib in NSCLC. Frontiers in immunology* 2019, 10:3135.
- 20.Kulkoyluoglu-Cotul E, Arca A, Madak-Erdogan Z: *Crosstalk between Estrogen Signaling and Breast Cancer Metabolism. Trends in endocrinology and metabolism: TEM* 2019, 30(1):25–38.



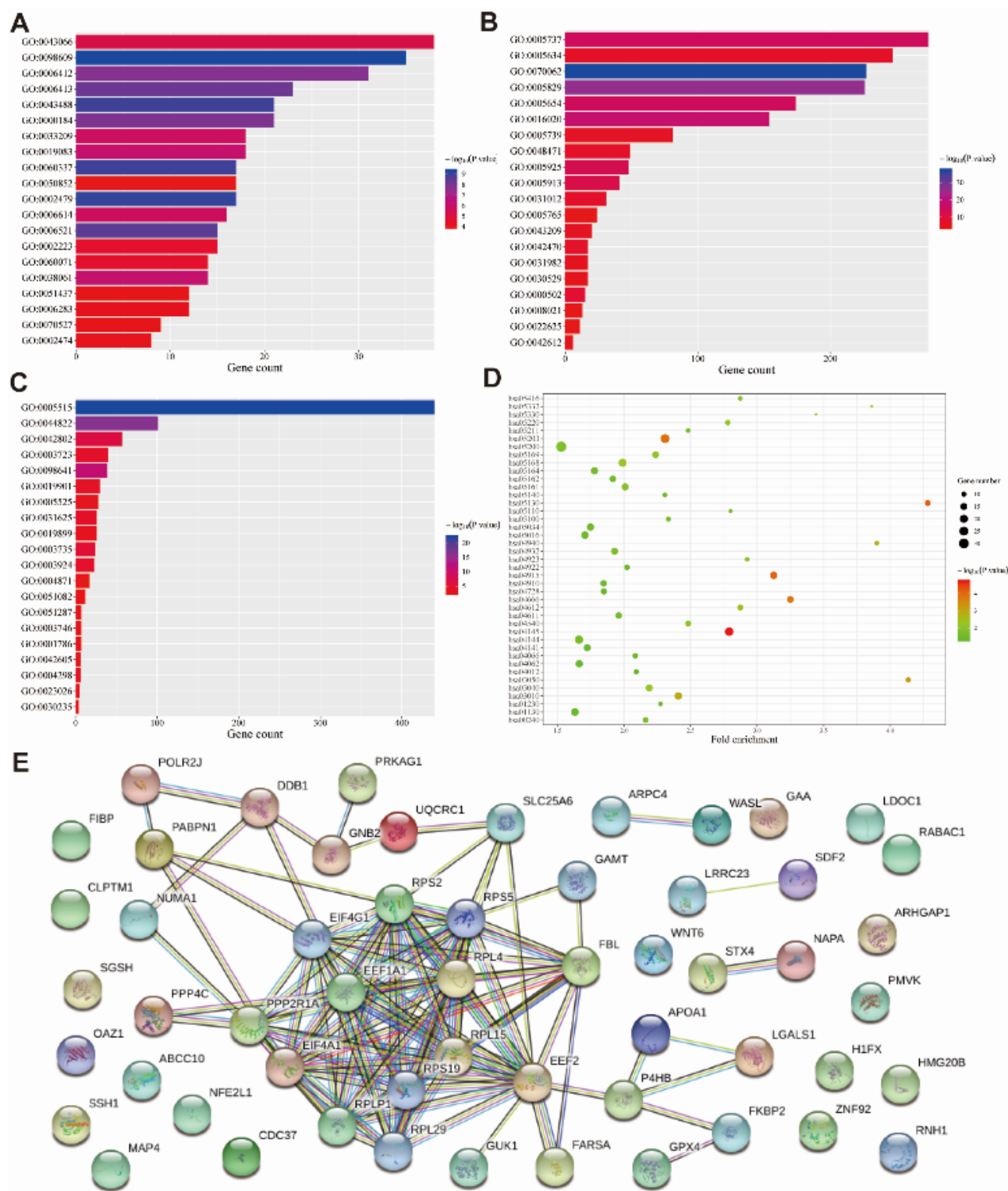
- 21.Piperigkou Z, Karamanos NK: *Estrogen receptor-mediated targeting of the extracellular matrix network in cancer. Seminars in cancer biology* 2020, 62:116–124.
- 22.Ebright RY, Lee S, Wittner BS, Niederhoffer KL, Nicholson BT, Bardia A, Truesdell S, Wiley DF, Wesley B, Li S *et al*: *Deregulation of ribosomal protein expression and translation promotes breast cancer metastasis. Science* 2020.
- 23.Wang M, Hu Y, Stearns ME: *RPS2: a novel therapeutic target in prostate cancer. Journal of experimental & clinical cancer research: CR* 2009, 28:6.
- 24.Wang M, Hu Y, Amatangelo MD, Stearns ME: *Role of ribosomal protein RPS2 in controlling let–7a expression in human prostate cancer. Molecular cancer research: MCR* 2011, 9(1):36–50.
- 25.Wang RX, Xu XE, Huang L, Chen S, Shao ZM: *eEF2 kinase mediated autophagy as a potential therapeutic target for paclitaxel-resistant triple-negative breast cancer. Annals of translational medicine* 2019, 7(23):783.
- 26.Tekedereli I, Alpay SN, Tavares CD, Cobanoglu ZE, Kaoud TS, Sahin I, Sood AK, Lopez-Berestein G, Dalby KN, Ozpolat B: *Targeted silencing of elongation factor 2 kinase suppresses growth and sensitizes tumors to doxorubicin in an orthotopic model of breast cancer. PloS one* 2012, 7(7):e41171.
- 27.Bayraktar R, Ivan C, Bayraktar E, Kanlikilicer P, Kabil NN, Kahraman N, Mokhlis HA, Karakas D, Rodriguez-Aguayo C, Arslan A *et al*: *Dual Suppressive Effect of miR–34a on the FOXM1/eEF2-Kinase Axis Regulates Triple-Negative Breast Cancer Growth and Invasion. Clinical cancer research: an official journal of the American Association for Cancer Research* 2018, 24(17):4225–4241.
- 28.Dong Z, Jiang H, Liang S, Wang Y, Jiang W, Zhu C: *Ribosomal Protein L15 is involved in Colon Carcinogenesis. International journal of medical sciences* 2019, 16(8):1132–1141.
- 29.He Z, Xu Q, Wang X, Wang J, Mu X, Cai Y, Qian Y, Shao W, Shao Z: *RPLP1 promotes tumor metastasis and is associated with a poor prognosis in triple-negative breast cancer patients. Cancer cell international* 2018, 18:170.

## Figures



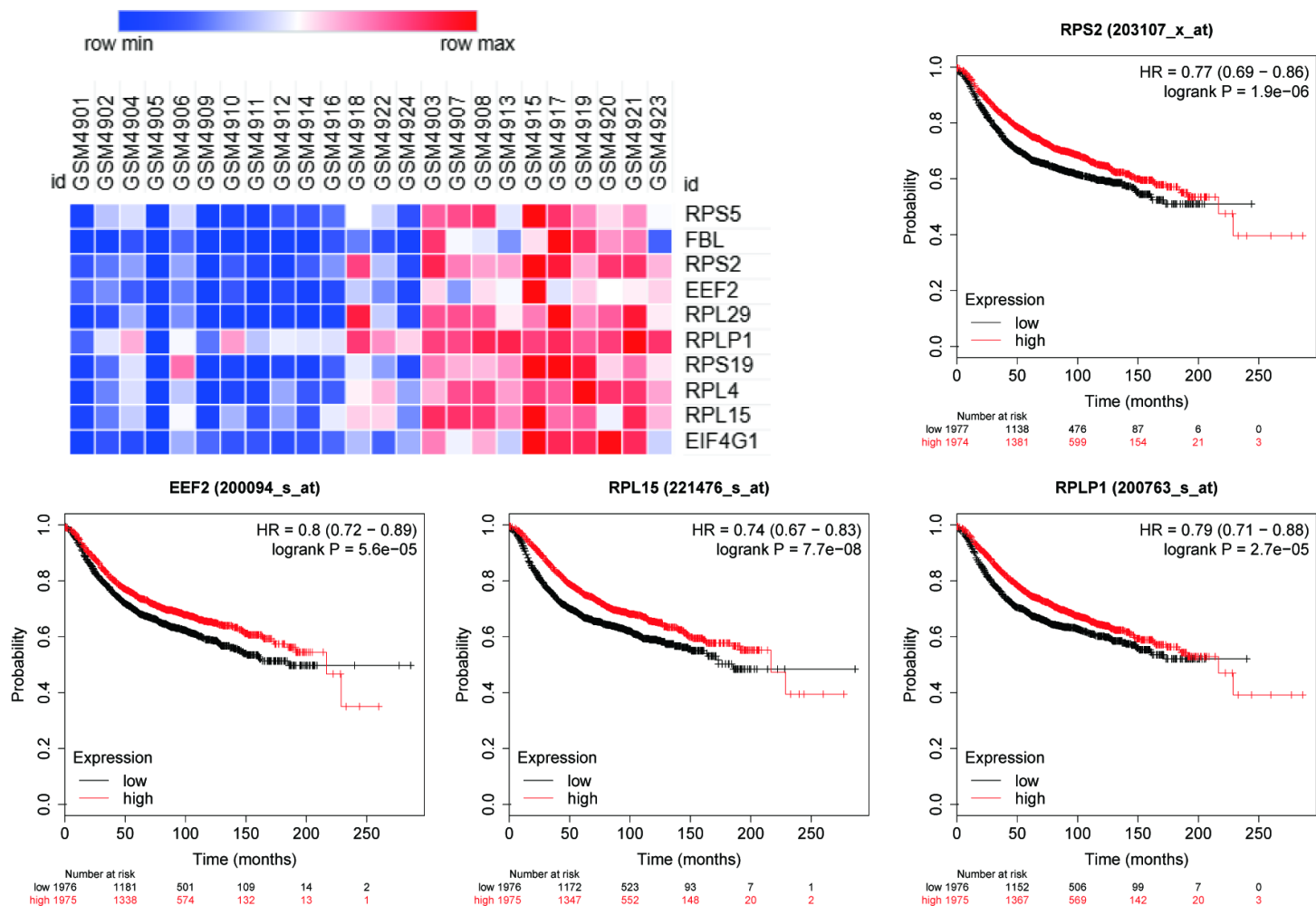
**Figure 1**

Construct of co-expression network and gene modules. A. The optimal soft threshold,  $\beta = 9$ . B. Distinguish 15 gene modules. C. Clinical information of breast cancer samples. D. Correlation between gene modules and clinical characters. E. Brown modules was most strongly associated with drug resistance ( $\text{cor}=0.85$ ,  $p < 1e-200$ ), including 735 genes.



**Figure 2**

KEGG-GO-PPI analysis. A. Top 20 in GO (Biological Process). B. Top 20 in GO (Cellular Component). C. Top 20 in GO (Molecular Function). D. Pathways in KEGG analysis (p value < 0.05). E. PPI network analysis with 55 hub genes.



**Figure 3**

10 Hub gene expression levels and their relationships to OS in BC patients. RPS2, EEF2, RPL15 and RPLP1 were significantly correlated with OS ( $P < 0.05$ ). The higher expression level indicated the better prognosis.

## Supplementary Files

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

- [Supplementary.docx](#)