Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Bioinformatic analysis and experimental validation identify prognostic hub genes associated with apigenin response in breast cancer

Zhengrong Zhang
Anhui university of chinese medicine
Fei Lu
Anhui University of Chinese Medicine
Hang Gao
Anhui university of chinese medicine
Yang Ge
Anhui university of chinese medicine
Yu Sheng
Anhui university of chinese medicine
Fengyuan Chen
Anhui university of chinese medicine
Hang Song
hangsong@ahtcm.edu.cn
Anhui university of chinese medicine

Research Article

Keywords: Breast cancer, Apigenin, Hub genes, WGCNA, Bioinformatics

Posted Date: November 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11049725/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Bioinformatic analysis and experimental validation identify prognostic hub genes associated with apigenin response in breast cancer

Zhengrong Zhang[#1]; Fei Lu[#2]; Hang Gao[1]; Yang Ge[1]; Yu Sheng[1]; Fengyuan Chen[*1]; Hang Song[*1,3,4]

1. School of Integrated Chinese and Western Medicine, Anhui University of Chinese Medicine, Longzihu Road 350, Hefei, China.
2. School of Acupuncture-Moxibustion and Tuina, Anhui University of Chinese Medicine, Longzihu Road 350, Hefei, China.
3. Functional Activity and Resource Utilization of Edible and Medicinal Fungi Anhui Provincial Joint Key Laboratory, Anhui University of Chinese Medicine, 350 Longzihu Road, Hefei, 230012, Anhui, P. R. China.
4. Anhui Provincial Key Laboratory of Tumor Evolution and Intelligent Diagnosis and Treatment, Bengbu, Anhui, P. R. China.

[#]These authors contributed equally to this work.

Zhengrong Zhang, Email: zrzhang311@126.com ORCID: https://orcid.org/0009-0001-0773-888X

Fei Lu, Email: lf20249090@stu.ahtcm.edu.cn ORCID: https://orcid.org/0009-0004-0669-3271

Hang Gao, Email: hanggao2024@163.com ORCID: https://orcid.org/0009-0007-7920-942X

Yang Ge, Emai: geyangcro@163.com ORCID: https://orcid.org/0009-0000-1768-1285

Yu Sheng, Emai: sy197693@163.com ORCID: https://orcid.org/ 0009-0007-5038-2229

***Correspondence to:**

Fengyuan Chen, Email: isobellachen@163.com ORCID: https://orcid.org/0009-0000-7769-6868

Hang Song, Email: hangsong@ahtcm.edu.cn ORCID: https://orcid.org/0000-0002-8154-2260

Abstract

Background: Breast cancer (BC) has emerged as the second most common cancer globally, making the exploration of its molecular mechanisms and novel therapeutic targets an urgent priority.

Objectives: This study aimed to identify key genes mediating the anti-BC effects of apigenin through bioinformatics analysis.

Methods: Weighted gene co-expression network analysis (WGCNA) was used to identify BC-associated gene modules (turquoise, yellow, and brown). Genes from these modules were intersected with differentially expressed genes (DEGs) from apigenin-treated BC cells, and then protein-protein interaction (PPI) analysis was performed to identify hub genes. The TCGA-BRCA dataset, the Gene Expression Profiling Interactive Analysis (GEPIA), and the Human Protein Atlas were employed for expression validation, as well as for prognostic and immune-infiltration analyses. KEGG enrichment analysis, molecular docking, and Western blotting were carried out to elucidate pathways, ligand-binding affinities, and the changes of protein-level expression.

Results: The intersection analysis and PPI screening jointly identified 10 hub genes. Validation using TCGA-BRCA, GEPIA, and the Human Protein Atlas confirmed that these genes exhibited distinct expression patterns between tumor and normal tissues, along with significant associations with survival outcomes and immune-cell infiltration. KEGG enrichment analysis suggested that these hub genes are involved in arginine and proline metabolism, as well as in cancer-related pathways. Molecular docking and Western blotting demonstrated that apigenin exhibits high binding affinity for IPO4, RRP9, and METTL1, and that the downregulation of their protein expression accompanies this.

Conclusion: The 10 hub genes could serve as potential prognostic biomarkers and candidate therapeutic targets for apigenin-based BC therapy, thereby providing mechanistic insights into the antitumor effects mediated by this traditional Chinese medicine (TCM) monomer.

Keywords: Breast cancer, Apigenin, Hub genes, WGCNA, Bioinformatics

1. Introduction

Breast cancer (BC) is one of the most common malignant tumors affecting women, arising primarily from breast tissue and typically manifesting with symptoms including breast lumps, skin changes, and nipple discharge. Although patient survival rates have improved modestly as a result of the continuous advances in early diagnosis and treatment of BC, both incidence and mortality continue to rise, particularly among middle-aged and elderly women [1]. In 2022, there were 2.3 million new cases and 670,000 female breast cancer deaths worldwide. In half of the countries surveyed, the annual rate increased by 1-5%. By 2050, the number of new cases and deaths is expected to increase by 38% and 68%, respectively [2]. Although radiotherapy and chemotherapy have made remarkable advances in BC treatment in recent years, most current drugs continue to exhibit considerable toxic side effects and drug resistance [3; 4]. Therefore, it is imperative to further identify risk factors that influence the occurrence and progression of BC and to develop therapeutic strategies.

Apigenin is a natural flavonoid compound with the chemical name 4',5,7-trihydroxyflavone. It exhibits diverse pharmacological effects, encompassing neuromodulation [5], hypoglycemic activity, anti-apoptotic properties, anti-inflammatory and antioxidant properties [6]. Previous studies have shown that the anti-cancer mechanisms of apigenin include suppression of cell migration and invasion, induction of cell-cycle arrest, promotion of apoptosis, and interference with tumor-associated signaling pathways [7; 8]. Apigenin has been shown to suppress highly metastatic human triple-negative BC by interfering with YAP/TAZ transcription activity [9]. Collectively, these findings indicate that apigenin possesses significant anti-tumor potential as a bioflavonoid. Nevertheless, systematic investigations into the mechanism of apigenin in BC are still limited, and its key molecular targets and regulatory networks remain to be elucidated. This knowledge gap severely limits its clinical translation and practical application.

In recent years, the rapid progress in bioinformatics has provided powerful tools for integrating multi-omics data and identifying key genes associated with diseases. Gene chips and microarrays have been widely used to screen for differentially expressed genes (DEGs) [10]. Weighted Gene Co-expression Network Analysis (WGCNA) identifies gene modules that are highly correlated with clinical features in high-throughput expression data, thereby facilitating the discovery of functional synergistic networks underlying disease pathogenesis. Furthermore, by integrating expression analysis,

protein interaction networks, and multi-database validation, hub genes that play pivotal roles in disease progression and drug response can be systematically identified.

On this basis, the present study aims to systematically identify potential molecular targets of apigenin in BC by integrating bioinformatics approaches with experimental validation, and to evaluate their clinical significance. WGCNA was employed to identify key gene modules from BC cell expression data. Intersection of these modules with the genes differentially expressed following apigenin intervention then yielded a set of candidate hub genes. Subsequently, multi-omics databases were used to validate their expression patterns, prognostic significance, and their associations with immune cell infiltration and DNA methylation. Finally, molecular docking and Western blot analysis were performed to verify the binding affinity and expression regulation of the key targets, thereby offering targeted therapeutic insights into the application of traditional Chinese medicine (TCM) in cancer treatment (Figure 1).

2. Materials and methods

2.1 Data preparation

The BC datasets GSE48213, GSE119552, and GSE33447 were obtained from NCBI Gene Expression Omnibus (GEO)[11]. GSE48213 contains high-throughput sequencing expression profiles from 56 BC cell lines, generated using the Illumina Genome Analyzer IIx platform. GSE119552 comprises gene expression data from MCF-7 breast cancer cells after 24 h treatment with E2, zearalenone, or apigenin. Each treatment condition was performed in four independent replicates. The expression profiles from apigenin-treated and untreated MCF-7 tumor cells were then screened and analyzed. The platform is Agilent-039494 SurePrint G3 Human GE v2 8x60K Microarray 039381. GSE33447 comprises 8 breast tumor tissues and 8 matched non-tumor breast tissues. This dataset was generated using the Agilent-028004 SurePrint G3 Human GE 8×60K Microarray. Following the download of normalized data, custom R scripts were utilized to map probe IDs to gene symbols. To construct a gene-level expression matrix, expression values from multiple probes mapping to the same gene were averaged. To mitigate potential batch effects arising from different data sources, the removeBatchEffect function from the limma package was applied. Survival differences across patient groups stratified by gene expression levels were evaluated using the Kaplan–Meier method and log-rank tests. Furthermore, statistically significant fold changes were visualized in a volcano plot

generated with the ggplot2 package (False discovery rate < 0.05 and |log2(fold-change)| > 1). All the aforementioned data processing and analyses were conducted within the R environment.

2.2 Construction of WGCNA

The expression data was preprocessed and normalized using R. The WGCNA R package was employed to construct the gene co-expression network. Initially, outliers in the RNA expression data were filtered out. Genes were sorted by their standard deviation (SD) values, and the top 5,000 genes were retained for WGCNA analysis. Soft-thresholding applies continuous element-wise thresholding to the adjacency matrix via a power function, and the selection of the β parameter affects the module detection results. To achieve a scale-free network topology, a β value of 9 was selected, yielding a scale-free fit index of $R^2 = 0.85$. The adjacency matrix was then transformed into a topological overlap matrix (TOM) to quantify network connectivity. Furthermore, average-linkage hierarchical clustering was performed based on the TOM dissimilarity measure, and the resulting dendrogram was further analyzed for module identification.

2.3 Identification of Clinically Significant Modules

Co-expression modules are highly topologically overlapping and share similar gene sets —genes within a module exhibit higher co-expression. Two key metrics were used to identif y modules significantly associated with clinical features. The module eigengene (ME) is defi ned as the first principal component of a given module and serves to summarize the express ion profile of that module across each sample. Module membership (MM) is the correlation coefficient between genes and module characteristic genes. Finally, the correlation coefficient s between each ME and the clinical traits were calculated to identify the modules most stron gly associated with the clinical features of interest.

2.4 Identification of key genes

Prior to analysis, probe IDs in each dataset were converted to corresponding gene symb ols. The Limma package was employed to identify DEGs from the expression matrix compari ng apigenin-treated with untreated MCF-7 tumor cells. Two thresholds were selected: False d iscovery rate < 0.05 and |log2 (fold-change)| > 1, and a volcano plot was generated accordin gly. Subsequently, based on the DEGs, hierarchical clustering analysis of their expression val ues was performed and visualized as a heatmap with centered Pearson correlation. Finally, th

e overlap between the DEGs and the significant gene modules was determined.

2.5 Function enrichment analysis

The Database for Annotation, Visualization and Integrated Discovery (DAVID)[12] is a publicly available bioinformatics resource that integrates diverse biological data and analytical tools to provide systematic and comprehensive functional annotation for large gene or protein lists. DAVID was used to perform functional annotation of the identified key genes. The enrichment results were visualized using R software.

2.6 Protein-protein interaction (PPI) network analysis and hub gene screening

The STRING database is a platform that predicts potential PPIs. These interactions encompass both direct physical binding between proteins and indirect functional associations[13]. Analyzing the interactions and functional relationships among key genes can provide valuable insights into disease pathogenesis and progression. The PPI network was constructed and visualized using Cytoscape (v. 3.5.0). It has been reported that protein expression levels are positively correlated with the functional significance of their corresponding genes. Thus, highly expressed genes are more likely to represent key functional genes. CytoHubba, a Cytoscape plugin, was used to identify hub nodes within the PPI network. The top 10 genes, ranked by their node degrees, were designated as hub genes.

2.7 Validation of hub genes

Gene Expression Profiling Interactive Analysis (GEPIA)[14] is an online database that enables analysis of RNA-sequencing expression data derived from 9,736 tumor and 8,587 normal samples in the TCGA and GTEx projects. The expression levels of the hub genes in TCGA-BRCA were visualized with the GEPIA platform. Protein-level expression of the hub genes was further validated using the Human Protein Atlas portal[15].

2.8 Association between hub genes and tumor-infiltrating immune cells

Tumor Immune Estimation Resource (TIMER)[16] is an online tool that utilizes RNA-seq expression data to estimate immune cell infiltration levels in tumor microenvironments. This tool was employed to investigate the correlations between the 10 hub genes and the abundance of tumor-infiltrating immune cells, including B cells, $CD4^+$ T cells, $CD8^+$ T cells, neutrophils, macrophages, and dendritic cells.

2.9 Methylation analyses of hub genes

The UALCAN portal is a comprehensive, user-friendly, and interactive web-based resource for the analysis of cancer multi-omics data. The platform enables the assessment of epigenetic gene regulation via promoter DNA methylation[17]. Specifically, the methylation levels of the hub genes were compared between tumor and normal tissues, and the correlation between their expression and corresponding methylation levels was evaluated using the methylation data available in the TCGA-BRCA dataset.

2.10 Genetical alteration of hub genes

The cBioPortal for Cancer Genomics is a comprehensive platform that facilitates the exploration and visualization of multidimensional cancer genomics data across genes, samples, and data types[18]. It supports cancer research by enabling the integration and interpretation of genetic, epigenetic, transcriptomic, and proteomic data. Through its interactive query interface, the platform provides comprehensive genomic profiles, allowing researchers to explore and compare genetic alterations across different samples.

2.11 Molecular Docking

The molecular structure of the active component apigenin was constructed using ChemDraw, saved in PDBQT format, and used as the docking ligand. Crystal structures of ten target proteins (including MRTO4 and BYSL) were obtained from the RCSB Protein Data Bank (PDB)[19]. The protein structures were prepared for docking by removing free water molecules, adding hydrogen atoms, and calculating Gasteiger charges. Molecular docking was performed using AutoDockTools (ADT) software. A docking grid encompassing the predicted active site of each protein was defined, followed by docking simulations that generated DLG result files. The 3D conformation exhibiting the highest binding affinity and the most stable intermolecular interactions was selected. This conformation was exported as a PDBQT file, and the docking results were visualized using PyMOL.

2.12 Cell culture and transfection

Human breast cancer cell lines MCF-7, MDA-MB-231 and SK-BR-3 were purchased from Procell (Wuhan, China; CL-0149, CL-0150B, CL-0211). According to the quality-control certificates provided by the supplier, all cell lines were authenticated by STR profiling and tested negative for mycoplasma contamination before delivery. MCF-7 and MDA-MB-231 cells were cultured in high-glucose DMEM medium (Procell, PM150210), while SK-BR-3 cells were maintained in McCoy's 5A medium (Procell, CM-0211). All culture media were supplemented with 10% fetal bovine

serum (FBS, HyClone, SV30087.02) and 1% penicillin-streptomycin solution (P/S, Biosharp, BL505A). Cells were cultured at 37 °C with 5% CO_2 and saturated humidity. Apigenin (Yuanye, A10033) was dissolved in DMSO (Biosharp, BL165B) to prepare a 100 mM stock solution. The stock solution was sterile-filtered through a DMSO-compatible 0.22 μm filter, aliquoted away from light, stored at −20 °C, and protected from repeated freeze-thaw cycles. Before experiments, apigenin was diluted to the working concentrations with complete culture medium.

Breast cancer cells (MCF-7, MDA-MB-231, SK-BR-3) in logarithmic growth phase were trypsinized and counted, then seeded into 6-well plates at a density of 2×10^5 cells per well. When cell confluence reached approximately 70–80%, plasmid transfection was performed using Lipofectamine™ 3000 transfection reagent (Thermo Fisher Scientific, L3000008) following the manufacturer's instructions. The full-length coding sequences of human IPO4, METTL1, and RRP9 were individually cloned into the pcDNA3.1 vector to construct overexpression plasmids pcDNA3.1-IPO4, pcDNA3.1-METTL1, and pcDNA3.1-RRP9, respectively. An empty pcDNA3.1 plasmid without target gene insertion was used as the negative control (Vector). Each overexpression plasmid or empty vector was transfected into the corresponding breast cancer cells. All plasmids were constructed and supplied by Yibeirui Biosciences (China). At 24 h post-transfection, the medium was replaced with fresh complete medium containing the designated concentration of apigenin, or with an equal volume of DMSO, for another 24 h of incubation. Cells were harvested afterwards for subsequent experiments.

2.13 CCK-8 assay for detection of cell proliferative viability

MCF-7, MDA-MB-231 and SK-BR-3 cells in logarithmic growth phase were trypsinized, counted and seeded into 96-well plates. MCF-7 and SK-BR-3 cells were seeded at a density of 5×10^3 cells per well, while MDA-MB-231 cells were seeded at 3×10^3 cells per well. Each well contained 100 μL of complete medium. After cell attachment, the culture medium was replaced with fresh complete medium containing serially diluted apigenin; an equal volume of DMSO was added to the vehicle control group. Following 24 h of apigenin treatment, 10 μL of CCK-8 reagent (Beyotime, C0037) was added to each well, and the plates were incubated at 37 °C for 2 h. The absorbance at 450 nm was subsequently measured using a microplate reader. Cell viability of each treatment group was normalized to the vehicle control group (defined as 100% viability). Based on the logarithmic values of apigenin concentrations on the X-axis, a four-parameter nonlinear regression model was used to fit

the dose–response curve and calculate the half-maximal inhibitory concentration (IC_{50}) in GraphPad Prism 10.6.1.

2.14 Western blot

Total protein was extracted from MCF-7, MDA-MB-231, and SK-BR-3 cells following treatment with cell line- and experiment-specific concentrations of apigenin. The protein concentration of cell lysates was determined using the BCA protein assay kit (Beyotime, P0010). After sample loading, proteins were separated via 10% SDS-PAGE electrophoresis and then transferred onto Polyvinylidene Fluoride (PVDF) membranes (Millipore, IPVH00010) by wet transfer. The membranes were blocked for 2 h and subsequently incubated with primary antibodies overnight at 4 °C. Primary antibodies used included RRP9 (Abcam, ab168849), IPO4 (Proteintech, 67549-1-Ig), METTL1 (Immunoway, YN5872), and the internal reference GAPDH (ProMab, YM8394). The dilution ratio was 1:1000 for RRP9, IPO4 and METTL1, and 1:10000 for GAPDH. After overnight incubation at 4 °C, membranes were incubated with goat anti-rabbit IgG secondary antibody (Immunoway, RS0002) at room temperature for 2 h. Protein bands were visualized by chemiluminescence, and signal intensities were analyzed using ImageJ software.

2.15 Statistical analysis

Survival analysis of the hub genes was performed using GEPIA, and those with statistically significant results (log-rank $P < 0.05$) were retained, highlighting the potential critical roles of these genes. The LinkedOmics database provides multi-omic and clinical data across 32 cancer types, comprising up to 11,158 patients from TCGA project[20]. The LinkFinder module allows flexible exploration of associations between selected clinical traits and all other available attributes within a given cancer cohort. Survival Analysis can be performed using all samples in a selected cohort, or a subset of the samples. The predictive performance of the hub genes was further evaluated via receiver operating characteristic (ROC) curve analysis on the GSE33447 dataset using R software.

Statistical analyses were performed using SPSS 27.0 software, and graphs were generated with GraphPad Prism 10.6.1 software. All *in vitro* experiments in this study were conducted with independent replicates. All data are expressed as mean ± standard deviation ($\bar{x} \pm$ s). Statistical differences among multiple groups were evaluated using one-way analysis of variance followed by Tukey's multiple-comparisons test. $P < 0.05$ was considered statistically significant.

3 Results

3.1 Construction of Weighted Gene Co-Expression Network

Sample clustering was first performed to identify potential outliers (Figure 2A). Soft threshold selection is an essential step in constructing a weighted gene co-expression network. The soft-thresholding power was evaluated over the range 1 to 20 to balance scale-free topology fit and average connectivity. A scale-free topology fit index of $R^2 = 0.85$ was achieved at a soft-thresholding power of $\beta = 9$ (Figure 2B–C). Using this power value, a hierarchical clustering tree (dendrogram) was constructed based on the top 5000 genes ranked by standard deviation, yielding 7 distinct co-expression modules. The seven modules contained 2808 (turquoise), 1231 (blue), 458 (brown), 168 (yellow), 151 (green), 122 (red), and 62 (gray) genes (Figure 2D). Genes that could not be assigned to any other module were grouped into the gray module, which was excluded from subsequent analyses. The interaction analysis among co-expressed genes across modules was presented in Figure 2E. The horizontal and vertical axes were color-coded to represent different modules. The intensity of red coloration in the heatmap reflects the connectivity strength between modules. No strong interactions were observed among modules, indicating that they are highly independent of one another.

3.2 Identification of Clinically Significant Modules

For each module, we summarize the expression of the gene co-expression eigengene (i.e., the module to which the first component belongs). The correlation between each module eigengene and BC molecular subtypes was then calculated. A heatmap was generated to visualize the associations between module eigengenes and BC molecular subtypes. The brown, yellow, and turquoise modules exhibit the highest correlation with BC molecular subtypes and are therefore selected as clinically significant modules (Figure 3A–D).

3.3 Identification of DEGs and functional enrichment

Differential expression analysis of GSE119552 identified a total of 319 DEGs, comprising 242 down-regulated and 77 up-regulated genes. A heatmap was generated to visualize the expression patterns of the DEGs (Figure 4A). The volcano plot displays the DEGs, with red and blue points indicating up-regulated and down-regulated genes, respectively (Figure 4B). A total of 101 genes overlapped between the three clinically significant modules and the DEGs from GSE119552 (Figure 4C). Gene Ontology (GO) enrichment analysis revealed that, under the cellular component (CC) category, the key genes were predominantly enriched in the nucleus, cytosol, nucleoplasm, anchored

component of plasma membrane, preribosome, large ribosomal subunit precursor, cytoplasm, membrane, MCM complex, and perinuclear region of cytoplasm (Figure 4E). For the molecular function (MF) category, the key genes were mainly enriched in protein binding, transcription factor binding, poly(A) RNA binding, identical protein binding, chromatin DNA binding, core promoter binding, protein dimerization activity, and neutral amino acid transmembrane transporter activity (Figure 4E). KEGG pathway analysis revealed that these key genes were primarily involved in arginine and proline metabolism, bladder cancer, and central carbon metabolism in cancer (Figure 4D). Additionally, the biological process (BP) category indicated enrichment in processes including amino acid transport, negative regulation of transcription, cellular response to hypoxia, rRNA processing, cellular response to epidermal growth factor stimulus, and negative regulation of the intrinsic apoptotic signaling pathway in response to hydrogen peroxide (Figure 4E).

3.4 Validation of hub genes

The expression levels of the hub genes differ between tumor and normal samples. To validate these findings, we further analyzed TCGA-BRCA data, which revealed that MYBBP1A was significantly downregulated in tumor tissues, whereas the remaining nine hub genes exhibited significantly higher expression (Figure 5A–J). In addition, immunohistochemistry (IHC) staining data from the Human Protein Atlas were used to assess the protein expression of the hub genes, with representative patient samples shown in Figure 6A–I. The IHC results were generally consistent with the transcriptional expression patterns. Collectively, these findings support the reliability of the hub gene identification strategy.

3.5 Survival Analysis and ROC Evaluation

Prognostic assessment is a critical aspect of tumor biology and has garnered extensive research interest. To evaluate the prognostic relevance of the 10 hub genes potential targeted by apigenin, survival analysis was performed (Figure 7A–J). Patients were stratified into high- and low-expression groups according to the median expression level of each hub gene, and the Cox proportional-hazards model was applied for survival analysis. Survival analysis revealed that the expression levels of these 10 genes were significantly associated with overall survival (OS) and patient survival duration. The association between these hub genes and BC prognosis suggests that they may participate in tumorigenesis and that their expression is potentially modifiable by apigenin. To further evaluate the potential of the hub genes as prognostic biomarkers for BC, receiver operating characteristic (ROC)

curves were generated, and the area under the curve (AUC) was calculated (Figure 8A–G). The results indicated that seven of the hub genes exhibited high predictive accuracy. Notably, MRTO4, BYSL, RRS1, NOP2, RRP9, METTL1, and IPO4 demonstrated superior predictive performance.

3.6 Effect of Hub gene expression on immune cells

Tumor-infiltrating immune cells play a critical role in the pathogenesis of BC. In this study, the TIMER online tool was employed to examine the correlations between the hub genes and tumor purity as well as immune cell composition in BC (Figure 9A–J). The analysis revealed that RRP9 was significantly correlated with tumor purity in BC. Regarding immune cell associations, NOP2 was correlated with neutrophils; RRS1 with $CD8^+$ T cells and macrophages; RRP9 with $CD8^+$ T cells and neutrophils; MYBBP1A with $CD8^+$ and $CD4^+$ T cells; EXOSC5 with $CD8^+$ T cells, macrophages, neutrophils, and dendritic cells (though weakly with $CD4^+$ T cells); and IPO4 with $CD8^+$ T cells and macrophages.

3.7 Relationship between hub gene expression and methylation state

To explore potential epigenetic mechanisms regulating gene expression in BC, the TCGA-BRCA methylation dataset was used to investigate the correlation between the expression levels of the 10 hub genes and their promoter methylation status (Figure 10). The results revealed that the methylation levels of MRTO4, BYSL, and EXOSC5 were significantly lower in BC tissues compared tonormal tissues, whereas RRS1, NOP2, RRP9, MYBBP1A, METTL1, NOP16, and IPO4 exhibited the opposite trend.

3.8 Genetic alteration of hub genes

The OncoPrint module of cBioPortal was used to display the genetic alteration status of the 10 hub genes in patients from the TCGA-BRCA cohort. Overall, 24.4% of the patients harbored genetic alterations in at least one of the 10 hub genes, with the alteration frequency for each gene detailed in Figure 11A. RRS1 and NOP2 exhibited the highest alteration frequencies (9% and 4%, respectively), with amplification and deep deletion being the predominant alteration types. Across BC subtypes, amplification remains the most frequent alteration type (Figure 11B). Figure 11C-J provides a detailed summary of all mutations: MRTO4 harbored 1 missense mutation; BYSL had 2 truncating mutations; RRS1 had 2 missense mutations and 1 in-frame mutation; NOP2 had 6 missense mutations; RRP9 had 1 missense, 2 truncating, and 1 in-frame mutation; MYBBP1A had 3 missense mutations; EXOSC5 had 2 missense mutations; and IPO4 had 4 missense mutations.

3.9 Molecular Docking Analysis

Molecular docking of apigenin with 10 target proteins was performed using the AutoDock Tool software. The calculated binding energies and hydrogen-bonding amino acid residues involved are summarized in Table 1. Lower binding energy indicates stronger binding affinity and greater docking stability. The receptor–ligand complexes were visualized using PyMOL to illustrate the optimal binding sites between apigenin and each protein receptor (Figure 12). Among them, the three target proteins, IPO4, METTL1 and RRP9, exhibited relatively low binding energies with apigenin, demonstrating the optimal binding affinities. Specifically, molecular docking results revealed that the binding energies of IPO4, METTL1, and RRP9 with apigenin were -8.0, -8.6, and -8.8 kcal/mol, respectively. Additionally, conformational analysis indicated that apigenin formed stable interaction networks with key amino acid residues within the binding pockets of these three target proteins (IPO4: ARG-765, LYS-276; METTL1: MET-142, SER-27; RRP9: SER-329, ALA-244, MET-407, SER-161). Based on the principle that a lower binding energy indicates a more stable ligand-receptor binding conformation, these three targets were selected for subsequent experimental validation.

3.10 Apigenin inhibits the viability and downregulates the expression of IPO4, METTL1 and RRP9 of BC cells in *vitro*

To investigate the effect of apigenin on the proliferative viability of breast cancer cells *in vitro*, human breast cancer cell lines MCF-7, MDA-MB-231 and SK-BR-3 were used in this study. As shown in Figure 13A–C, after 24 h of treatment with serially diluted apigenin, the viability of MCF-7, MDA-MB-231 and SK-BR-3 cells decreased gradually in a dose-dependent manner compared with the control group ($P<0.05$). Dose–response curves were plotted based on cell viability data to calculate the IC_{50} (Figure 13D–F). The IC_{50} value of apigenin for MCF-7, MDA-MB-231, and SK-BR-3 cells was 18.27 μM, 74.26 μM and 29.34 μM, respectively (Figure 13D–F). These results indicated that apigenin exerted prominent inhibitory effects on the growth of breast cancer cells.

To further clarify the regulatory effect of apigenin on the protein expression of IPO4, METTL1 and RRP9, gradient concentrations of apigenin were set according to the above IC_{50} results: 2.5, 5 and 10 μM for MCF-7 cells, 20, 40 and 60 μM for MDA-MB-231 cells, and 10, 15 and 20 μM for SK-BR-3 cells (Figure 13G–I). Western blot quantification results (Figure 13J–L) revealed that apigenin treatment dose-dependently suppressed the protein levels of IPO4, METTL1, and RRP9 in MCF-7, MDA-MB-231, and SK-BR-3 cells, with statistically significant differences compared with the control

group (P<0.05, P<0.01, P<0.001). Collectively, IPO4, METTL1 and RRP9 may serve as potential molecular targets mediating the anti-tumor activity of apigenin.

3.11 Apigenin inhibits the expression of key gene-related proteins in BC

To verify that apigenin exerts its anti-tumor activity by targeting IPO4, METTL1, and RRP9, overexpression plasmids for IPO4, METTL1, and RRP9 were constructed for rescue assays, and the experimental design is shown in Figure 14A. As shown in Figure 14B–G, apigenin treatment markedly downregulated the protein expression levels of IPO4, METTL1, and RRP9 across all three breast cancer cell lines compared with the control group (P<0.05). Compared with the Apigenin group, the protein abundance of IPO4, METTL1, and RRP9 was significantly restored in the OE+Apigenin group, effectively reversing the inhibitory effect induced by apigenin (P<0.05). In addition, the relative protein expression was slightly decreased in the OE+Apigenin group when compared with the OE group. Collectively, these findings demonstrate that apigenin inhibits breast cancer progression via regulating IPO4, METTL1 and RRP9.

4. Discussion

BC, a malignant tumor arising in breast tissue, is one of the most diagnosed malignancies in women worldwide [21]. Its substantial tumor heterogeneity [22], drug resistance during advanced-stage treatment [23], and intricate molecular regulatory mechanisms [24] remain major challenges for clinical diagnosis and treatment. Apigenin is a natural active constituent derived from TCM. It exhibits low toxicity and multi-target regulatory biological properties and has demonstrated promising application potential in anti-tumor research [25]. Nevertheless, the precise molecular mechanisms by which apigenin interferes with the initiation and progression of breast cancer have not yet been fully elucidated.

In this study, a gene co-expression network was constructed from the GSE48213 dataset, and three clinically relevant modules (Turquoise, Yellow, and Brown) were identified as strongly associated with breast cancer subtypes. We narrowed down the range of candidate targets by performing intersection analysis between DEGs in apigenin-treated breast cancer cells and hub genes within these modules. PPI network analysis screened out ten core genes (MRTO4, BYSL, RRS1, NOP2, RRP9, MYBBP1A, EXOSC5, METTL1, NOP16 and IPO4). Meanwhile, the biological functions of hub genes were explored from three perspectives: metabolic reprogramming, epigenetic modification, and tumor immune microenvironment. Finally, IPO4, METTL1 and RRP9 were recognized as the core potential

targets mediating the anti-tumor activity of apigenin. This strategy effectively integrates the biological effects of apigenin and the molecular characteristics of breast cancer, laying a foundation for exploring the mechanism of action of this monomer.

Apigenin is a naturally occurring flavonoid widely distributed in fruits and vegetables, which exerts potential anticancer effects against various malignancies[25], and it can inhibit the invasion of breast cancer cells by up to 92% [26]. BC cells are highly dependent on aberrant transcriptional programs and protein synthesis. Functional enrichment analysis in the present study revealed that these ten core genes are mainly involved in arginine/proline metabolism, cancer-associated signaling pathways, ribosome biogenesis, and RNA processing. Existing studies have demonstrated that arginine transfer between cancer cells and macrophages dominates arginine-driven breast cancer progression. This process activates the Peroxisome proliferator-activated receptor-γ (PPARG) pathway via TDG/p53-dependent DNA demethylation, promotes the immunosuppressive polarization of tumor-associated macrophages, thereby dampening the anti-tumor activity of $CD8^+$ T cells and accelerating the progression of breast cancer cells [27; 28]. Previous research has indicated that apigenin can directly target spermine oxidase (SMOX), decrease the spermidine/spermine (Spd/Spm) ratio, and suppress xenograft tumor growth[29]. Accordingly, whether apigenin regulates polyamine catabolism through arginine/proline metabolism warrants further investigation.

To further characterize the expression profiles of the ten hub genes in breast tumors, we integrated the TCGA-BRCA dataset, GEPIA database and Human Protein Atlas to conduct cross-validation at the transcriptional and protein levels. The results revealed significant differential expression of all hub genes between breast cancer and normal breast tissues with distinct expression patterns. The tumor suppressor MYBBP1A was markedly downregulated in cancer tissues, whereas the other nine hub genes associated with ribosome function, RNA modification and nuclear transport exhibited tumor-specific overexpression. MYBBP1A (Myb-binding protein 1a, also known as p160) is a multifunctional protein predominantly localized in the nucleolus. Accumulating evidence indicates that MYBBP1A expression is negatively correlated with breast tumorigenesis, and MYBBP1A suppresses breast tumorigenesis by reinforcing p53-dependent inactivation [30]. Moreover, apigenin can effectively upregulate p53 expression in breast cancer stem cells [31]. In addition, apigenin induces G0/G1 cell cycle arrest by activating tumor protein p53 in vascular smooth muscle cells (VSMCs) [32]. Collectively, these findings suggest that MYBBP1A serves as an important potential target through

which apigenin regulates the growth of breast cancer cells.

Among the remaining upregulated hub genes, molecular docking with apigenin revealed that IPO4, METTL1 and RRP9 exhibited relatively lower binding energy to apigenin compared with other hub genes and formed stable interaction complexes. The oncogenic functions of RRP9, METTL1 and IPO4 have been validated in previous studies. RRP9 is highly expressed in breast cancer tissues. Using the PrimeView Human Gene Expression Array and IPA analysis, JUN was identified as a potential downstream target of RRP9, which facilitates breast cancer progression via the AKT signaling pathway[33]. In the present study, Western blot assays performed on MCF-7, MDA-MB-231, and SK-BR-3 cell lines further confirmed the aberrant expression of RRP9, METTL1, and IPO4 in apigenin treated breast cancer cells. Moreover, rescue assays suggested apigenin inhibits breast cancer cell proliferation by targeting RRP9, METTL1, and IPO4. Collectively, our multi-omics validation results reaffirm the critical regulatory roles of these genes in breast cancer progression and support their potential as therapeutic targets for apigenin against BC.

On this basis, survival analysis and ROC curve evaluation were further performed to explore the clinical prognostic value of these hub genes. Kaplan–Meier survival curves revealed that the expression levels of all 10 hub genes were significantly correlated with overall survival in breast cancer patients. Among them, seven genes, including MRTO4, BYSL, RRS1, NOP2, RRP9, METTL1 and IPO4, exhibited reliable prognostic discriminatory power; high expression of these genes was associated with poorer survival outcomes in patients. These findings indicate that this set of genes is not only involved in the molecular pathogenesis of breast cancer but can also serve as stable prognostic biomarkers. Previous multi-gene signature studies on breast cancer have suggested that compared with single molecular biomarkers, multi-gene signature models can reduce prediction bias caused by tumor molecular subtypes and clinical stages and markedly improve the accuracy of prognostic evaluation [34; 35]. The seven genes with high prognostic efficacy identified in this study are strong candidates for constructing an apigenin-related risk prediction model for breast cancer. Meanwhile, TIMER immune infiltration correlation analysis demonstrated that the expression levels of multiple hub genes were significantly associated with the abundance of intratumoral immune cells. For instance, RRP9 expression was correlated with the infiltration of $CD8^+$ effector T cells, and NOP2 expression was linked to tumor-associated neutrophil infiltration. This suggests that this ribosome-regulatory gene network participates in shaping the tumor immune microenvironment via multiple pathways.

The tumor immune microenvironment plays a pivotal role in breast cancer progression and therapeutic response [36]. Accordingly, apigenin modulates the tumor immune microenvironment by regulating the expression of core genes, which offers a novel research direction for investigating its immunomodulatory effects. Studies have demonstrated that apigenin possesses favorable binding affinity with proteins associated with ferroptosis and glycolysis. *In vivo* and *in vitro* experiments further verified that apigenin suppresses tumor growth in 4T1 xenograft mice by modulating the expression of ferroptosis- and glycolysis-related proteins [37]. Moreover, apigenin can regulate alternative splicing and reprogram cancer-associated spliced isoforms (ASI) into isoforms detected in normal non-tumor tissues. Specifically, it affects the splicing of anti-apoptotic genes, attenuates tumor proliferation, and thereby restrains triple-negative breast cancer (TNBC) growth *in vivo* [8]. Consistent with these reports, our study also confirmed that apigenin exerts potent inhibitory effects on MCF-7, MDA-MB-231 and SK-BR-3 cells. Collectively, this evidence indicates that apigenin exerts anti-breast cancer effects via multiple pathways. Nevertheless, the molecular targets and functional network of apigenin in breast cancer remain poorly defined, hindering its clinical translation and application.

Aberrant activation of ribosome biogenesis represents an early event during breast tumorigenesis and persists throughout tumor progression[38]. Among the core genes screened in this study, MRTO4, RRS1, NOP2, RRP9, EXOSC5, METTL1, NOP16 and IPO4 are all ribosome-related genes. Sustained hyperactivation of ribosomal pathways in tumor cells markedly accelerates global protein synthesis, providing material support for tumor proliferation and invasion. Accumulating evidence reveals that c-MYC-mediated ribosome biogenesis confers resistance to docetaxel chemotherapy in TNBC [39], further highlighting that ribosomal pathways act as critical drivers of malignant progression and drug resistance in breast cancer. *In vitro* studies have reported that apigenin can directly target ribosomal protein S9 to downregulate cyclin-dependent kinase 1 (CDK1), arresting the growth of human colon cancer cells at the G2/M phase [40]. Additionally, ribosomal protein RPS9 is aberrantly overexpressed in TNBC and closely associated with malignant tumor phenotypes [41]. Combined with the pathway enrichment results of the present study, we reasonably hypothesize that apigenin suppresses breast cancer cell proliferation by targeting ribosome-related core genes. Collectively, all enrichment findings suggest that the anti-tumor activity of apigenin relies on disrupting essential biological programs in breast cancer cells, including ribosome biogenesis and RNA modification.

This study identified key genes with significant dysregulation in DNA methylation. Methylation

analysis revealed that the core genes exhibit distinct epigenetic regulatory patterns in BC. MRTO4, BYSL and EXOSC5 display promoter hypomethylation in BC tissues, whereas RRS1, NOP2 and RRP9 are characterized by hypermethylation. DNA methylation serves as a crucial epigenetic mechanism in the initiation and progression of breast cancer, facilitating tumor development by modulating the expression of pivotal genes [42]. Existing studies have demonstrated that apigenin not only enhances DNA methylation of ubiquitin-like TNF-α promoter [43], but also regulates DNA and histone methylation associated with epigenetic modification [44]. Genomic alteration analysis indicated that 24.4% of breast cancer patients harbor mutations in these ten core genes. Among them, RRS1 and NOP2 possess the highest mutation frequencies (9% and 4%, respectively). Amplification and deep deletion constitute the predominant variant types, suggesting that genetic aberrations may drive the dysregulation of core genes in breast cancer. Current research regarding the effects of apigenin on DNA methylation remains limited. Collectively, our findings imply that the epigenetic effects of apigenin in breast cancer merit further exploration.

Based on the results of molecular docking, conformational stability, biological functional correlation and clinical prognostic significance, IPO4, METTL1 and RRP9 were selected as priority candidates for experimental validation in the present study. Nevertheless, we acknowledge that other identified hub genes may also exert prognostic effects. Future studies are warranted to further validate these findings by integrating comprehensive clinical data from large-scale cohorts.

The present study demonstrated that apigenin inhibited breast cancer cell viability and downregulated the protein expression of IPO4, METTL1, and RRP9 *in vitro*; nevertheless, several notable limitations should be acknowledged. Previous pharmacological studies have indicated that apigenin has pharmacokinetic limitations, including poor water solubility [45] and low oral bioavailability [46]. Accordingly, *in vitro* cellular data alone cannot fully reflect its *in vivo* exposure, metabolism, effective dosage, or antitumor activity. Given the complexity of the *in vivo* tumor microenvironment, further animal studies are required to determine whether apigenin can achieve sufficient tumor exposure and exert sustained antitumor effects under physiological conditions. Future studies should also systematically evaluate its biodistribution, pharmacokinetic profile, effective dosage, and systemic safety, including the maximum tolerated dose, therapeutic index, and potential long-term toxicity. Such investigations will be essential to validate the proposed molecular mechanisms and assess the translational potential of apigenin in breast cancer treatment.

5. Conclusion

This study utilized the GSE48213 dataset to construct a gene co-expression network, identifying three clinically relevant modules (Turquoise, Yellow, Brown) highly associated with BC subtypes. By intersecting the DEGs from apigenin-treated BC cells with key genes in these modules, we narrowed the candidate target pool. Further PPI network analysis identified 10 core genes (MRTO4, BYSL, RRS1, NOP2, RRP9, MYBBP1A, EXOSC5, METTL1, NOP16, and IPO4). This strategy effectively integrates the biological effects of apigenin with the molecular characteristics of BC, laying the foundation for exploring the mechanism of action of this monomer.

Statements & Declarations

Funding

This work was supported by the Training Action Project of Anhui Provincial Education Department of Young and Middle-aged Teachers (grant no. JNFX2025022), Anhui University of Chinese Medicine "Basic-Clinical Integration" Special Program (grant no. JCLCA2025001), the Open Project of Anhui Provincial Key Laboratory of Tumor Evolution and Intelligent Diagnosis and Treatment (grant no. KFKT202504), the National Natural Science Foundation of China (grant no.82303222).

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Authorship contribution

ZZ and FL: Conceptualization, Writing – original draft. HG: Investigation, Visualization; YG and YS: Investigation, Methodology. FC and HS: Supervision, Resources, Writing– review & editing. The authors declare that all data were generated in-house, that no paper mill was used and that no AI tool has been used for the generation of text or figures.

Acknowledgments

During the preparation of this work, the authors used “deepseek and Grammarly” to improve the language and readability. After using this tool, the authors reviewed and edited the content as needed

and take full responsibility for the content of the published article."

Data Availability Statement

The public datasets used for the bioinformatics analyses in this study are available from the following repositories and resources. Transcriptomic datasets were obtained from the NCBI Gene Expression Omnibus (GEO) under accession numbers GSE48213, GSE119552, and GSE33447. Gene expression data from The Cancer Genome Atlas Breast Invasive Carcinoma project (TCGA-BRCA) and the Genotype-Tissue Expression (GTEx) project were accessed through GEPIA. Protein expression and immunohistochemical data were obtained from the Human Protein Atlas. Functional annotation and enrichment analyses were performed using DAVID, and protein–protein interaction data were obtained from STRING. Tumor-infiltrating immune cell estimates and promoter DNA methylation data for TCGA-BRCA were accessed through TIMER and UALCAN, respectively. Genomic alteration data were accessed through cBioPortal, while multi-omics and clinical association analyses used data available through LinkedOmics. Protein structures used for molecular docking were retrieved from the RCSB Protein Data Bank. The processed data and raw experimental data generated in this study are available from the corresponding author upon reasonable request.

Clinical trial number: not applicable

Ethics, Consent to Participate, and Consent to Publish declarations: not applicable.

References

1. Lei S, Zheng R, Zhang S, Wang S, Chen R, Sun K, Zeng H, Zhou J, Wei W. Global patterns of breast cancer incidence and mortality: A population-based cancer registry data analysis from 2000 to 2020. Cancer Commun (Lond). 2021;41:1183-94. https://doi.org/10.1002/cac2.12207.

2. Kim J, Harper A, McCormack V, Sung H, Houssami N, Morgan E, Mutebi M, Garvey G, Soerjomataram I, Fidler-Benaoudia M M. Global patterns and trends in breast cancer incidence and mortality across 185 countries. Nat Med. 2025;31:1154-62. https://doi.org/10.1038/s41591-025-03502-3.

3. Bi T, Zhao Q, Wang T, Huang R, Liu B, Liu X, Wang Y, Sun Q, Yang Y, Liu Z. Disruption of Ferroptosis Inhibition and Immune Evasion with Tumor-Activatable Prodrug for Boosted Photodynamic/Chemotherapy Eradication of Drug-Resistant Tumors. Adv Healthc Mater. 2025;14:e2403473. https://doi.org/10.1002/adhm.202403473.

4. Liu W, Xu J, Chen L, Zhang D, Zhang J, Lu L, Zhang X, Huang X, Zhang G. Discovery of a novel BCL-2 inhibitor GW806742X for the treatment of TNBC. Biochem Pharmacol. 2025;237:116925. https://doi.org/10.1016/j.bcp.2025.116925.

5. Wu JW, Zhou YT, Wang BX, Shen LP, Zhang XQ, Du ZY, Wang P, Lu XJ, Miao ZL, Zhao XD. Apigenin regulates CCR5/JAK1/STAT1/MMPs signaling to alleviate secondary brain injury after intracerebral hemorrhage and its enhanced delivery via targeted nanoparticles. J Nanobiotechnology. 2025;23:696. https://doi.org/10.1186/s12951-025-03748-6.

6. Kong X, Liu Y, Zhou Z, Zhu L, Ai X, Tang J, Guo J, Tian P, Chen X. Integrated network pharmacology, molecular docking, and experimental validation elucidate the anti-inflammatory and antioxidant mechanisms of apigenin in LPS-induced acute lung injury. RSC Adv. 2026;16:42679-96. https://doi.org/10.1039/d6ra05523k.

7. Mao M, Jia Y, Chen Y, Yang J, Xu L, Zhang X, Zhou J, Li Z, Chen C, Ju S, Wang L. HJURP regulates cell proliferation and chemo-resistance via YAP1/NDRG1 transcriptional axis in triple-negative breast cancer. Cell Death Dis. 2022;13:396. https://doi.org/10.1038/s41419-022-04833-6.

8. Sudhakaran M, Navarrete TG, Mejía-Guerra K, Mukundi E, Eubank TD, Grotewold E, Arango D, Doseff A I. Transcriptome reprogramming through alternative splicing triggered by apigenin drives cell death in triple-negative breast cancer. Cell Death Dis. 2023;14:824. https://doi.org/10.1038/s41419-023-06342-6.

9. Jung O, Baek MJ, Wooldrik C, Johnson KR, Fisher KW, Lou J, Ricks TJ, Wen T, Best MD, Cryns VL et al. Nuclear phosphoinositide signaling promotes YAP/TAZ-TEAD transcriptional activity in breast cancer. Embo j. 2024;43:1740-69. https://doi.org/10.1038/s44318-024-00085-6.

10. Rossi C, Cicalini I, Cufaro MC, Consalvo A, Upadhyaya P, Sala G, Antonucci I, Del Boccio P, Stuppia L, De Laurenzi V. Breast cancer in the era of integrating "Omics" approaches. Oncogenesis. 2022;11:17. https://doi.org/10.1038/s41389-022-00393-8.

11. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30:207-10. https://doi.org/10.1093/nar/30.1.207.

12. Sherman B T, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, Imamichi T, Chang W. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res. 2022;50:W216-w21. https://doi.org/10.1093/nar/gkac194.

13. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S et al. The STRING database in 2023: protein-protein association networks

and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51:D638-d46. https://doi.org/10.1093/nar/gkac1000.
14. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017;45:W98-w102. https://doi.org/10.1093/nar/gkx247.
15. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson A, Kampf C, Sjostedt E, Asplund A et al. Proteomics. Tissue-based map of the human proteome. Science. 2015;347:1260419. https://doi.org/10.1126/science.1260419.
16. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu J S, Li B, Liu XS. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. Cancer Res. 2017;77:e108-e10. https://doi.org/10.1158/0008-5472.Can-17-0307.
17. Chandrashekar DS, Karthikeyan SK, Korla PK, Patel H, Shovon AR, Athar M, Netto GJ, Qin ZS, Kumar S, Manne U et al. UALCAN: An update to the integrated cancer data analysis platform. Neoplasia. 2022;25:18-27. https://doi.org/10.1016/j.neo.2022.01.001.
18. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer M L, Larsson E et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2:401-4. https://doi.org/10.1158/2159-8290.Cd-12-0095.
19. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res. 2000;28:235-42. https://doi.org/10.1093/nar/28.1.235.
20. Vasaikar SV, Straub P, Wang J, Zhang B. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. Nucleic Acids Res. 2018;46:D956-d63. https://doi.org/10.1093/nar/gkx1090.
21. Xu Y, Gong M, Wang Y, Yang Y, Liu S, Zeng Q. Global trends and forecasts of breast cancer incidence and deaths. Sci Data. 2023;10:334. https://doi.org/10.1038/s41597-023-02253-5.
22. Wang M, Zhang Y, Li Z, Fu L. Reprogramming of amino acid metabolism in breast cancer. Biochim Biophys Acta Rev Cancer. 2025;1880:189412. https://doi.org/10.1016/j.bbcan.2025.189412.
23. Barraud S, Vacher S, Schiffler C, Wong J, Hamza A, Courtois L, Pinton A, Servant N, Baulande S, Attignon V et al. Genomic landscape of estrogen receptor-positive HER2-negative advanced breast cancer with acquired resistance to aromatase inhibitors: Identification of an ESR1 alteration-related gene signature. Br J Cancer. 2025;133:464-72. https://doi.org/10.1038/s41416-025-03083-5.
24. Chai Y, Shi Y. The role of genetics and epigenetics in breast cancer: A comprehensive review of metastasis, risk factors, and future perspectives. J Pharm Anal. 2025;15:101268. https://doi.org/10.1016/j.jpha.2025.101268.
25. Singh D, Gupta M, Sarwat M, Siddique HR. Apigenin in cancer prevention and therapy: A systematic review and meta-analysis of animal models. Crit Rev Oncol Hematol. 2022;176:103751. https://doi.org/10.1016/j.critrevonc.2022.103751.
26. Yuan L, Cao Y, Zhang Q, Pan J, Wu C, Ye Y, Jiao Q, Zhu HL, Wang Z. Rational design of mitochondria-targeted fluorescent biosensors for in vivo elucidation of the interaction between breast cancer metastasis and mitochondrial autophagy. Biosens Bioelectron. 2024;251:116123. https://doi.org/10.1016/j.bios.2024.116123.
27. Vogel I, Bapat SP. TAMing immunity through an unexpected source. Cancer Cell. 2025;43:1002-04. https://doi.org/10.1016/j.ccell.2025.04.004.
28. Zhu Y, Zhou Z, Du X, Lin X, Liang ZM, Chen S, Sun Y, Wang Y, Na Z, Wu Z et al. Cancer cell-derived arginine fuels polyamine biosynthesis in tumor-associated macrophages to promote immune evasion. Cancer Cell. 2025;43:1045-60.e7. https://doi.org/10.1016/j.ccell.2025.03.015.

29. Zhang Z, Xiang B, Geng R, Ji X, Wang D, Yin Z, Chen L, Liu J. Targeting Polyamine Metabolism in Colorectal Cancer: Apigenin Dismantles the HIF-1α/SMOX Positive Feedback Loop to Suppress Tumor Progression. Int J Mol Sci. 2026;27:3261. https://doi.org/10.3390/ijms27073261.
30. Akaogi K, Ono W, Hayashi Y, Kishimoto H, Yanagisawa J. MYBBP1A suppresses breast cancer tumorigenesis by enhancing the p53 dependent anoikis. BMC Cancer. 2013;13:65. https://doi.org/10.1186/1471-2407-13-65.
31. Ou M, Deng Z, Shi Y, He J, Ye Z, Guo M, Cheng G, Wu J, Lv L. Mechanism of Apigenin against breast cancer stem cells: network pharmacology and experimental validation. Front Pharmacol. 2024;15:1496664. https://doi.org/10.3389/fphar.2024.1496664.
32. Gao J, Wang D, Zhang X, Yang G, Xi D, Qin X, Wang Y, Jin Y, Guo Y, Li X, Ma K. Apigenin prevents hypertensive vascular remodeling by regulating the TP53 pathway. Int Immunopharmacol. 2025;157:114706. https://doi.org/10.1016/j.intimp.2025.114706.
33. Huan J, Liu X, Wang N, Mu Y, Li L, Du Y. The RRP9-JUN axis promotes breast cancer progression via the AKT signalling pathway. Biology direct. 2024;19:131. https://doi.org/10.1186/s13062-024-00578-8.
34. Guo X, Cao Y, Shi X, Xing J, Feng C, Wang T. Evaluating the prognostic potential of telomerase signature in breast cancer through advanced machine learning model. Front Immunol. 2024;15:1462953. https://doi.org/10.3389/fimmu.2024.1462953.
35. Zhang X, Min S, Zhang N, Shi X, Cai Z, Yang D, Meng Z, Zhao Y, Ni N, Wang T. Machine learning on transcription factor expression profiles for precision breast cancer therapy. Cancer Cell Int. 2025;25:346. https://doi.org/10.1186/s12935-025-03987-5.
36. Akinsipe T, Mohamedelhassan R, Akinpelu A, Pondugula SR, Mistriotis P, Avila LA, Suryawanshi A. Cellular interactions in tumor microenvironment during breast cancer progression: new frontiers and implications for novel therapeutics. Front Immunol. 2024;15:1302587. https://doi.org/10.3389/fimmu.2024.1302587.
37. Zhu M, Peng D, Yu Z, Zhao M, Liu J, Yu W, Xu H, Lu B, Liao N, Niu S et al. Apigenin Inhibits Triple-Negative Breast Cancer Growth by Dual Targeting GPX4/SLC7A11-Mediated Ferroptosis and PKM2/GLUT1-Mediated Aerobic Glycolysis. Phytother Res. 2026;40:3741-57. https://doi.org/10.1002/ptr.70337.
38. Metge BJ, Williams L, Swain C A, Hinshaw D C, Elhamamsy AR, Chen D, Samant RS, Shevde LA. Ribosomal RNA Biosynthesis Functionally Programs Tumor-Associated Macrophages to Support Breast Cancer Progression. Cancer Res. 2025;85:1459-78. https://doi.org/10.1158/0008-5472.Can-24-0707.
39. Wang X, Wang X, Gu J, Wei Y, Wang Y. circUBR5 promotes ribosome biogenesis and induces docetaxel resistance in triple-negative breast cancer cell lines via the miR-340-5p/CMTM6/c-MYC axis. Neoplasia. 2025;59:101062. https://doi.org/10.1016/j.neo.2024.101062.
40. Iizumi Y, Oishi M, Taniguchi T, Goi W, Sowa Y, Sakai T. The flavonoid apigenin downregulates CDK1 by directly targeting ribosomal protein S9. PLoS One. 2013;8:e73219. https://doi.org/10.1371/journal.pone.0073219.
41. Lin Z, Peng R, Sun Y, Zhang L, Zhang Z. Identification of ribosomal protein family in triple-negative breast cancer by bioinformatics analysis. Biosci Rep. 2021;41:BSR20200869. https://doi.org/10.1042/BSR20200869.
42. Kundu S, Sarkar S, Ghosh S, Acharya Chowdhury A. DNA methylation as an oncogenic driver in breast cancer: Therapeutic targeting via epigenetic reprogramming of DNA methyltransferases.

Biochem Pharmacol. 2025;242:117313. https://doi.org/10.1016/j.bcp.2025.117313.
43. Mei X, Wang J, Zhang C, Zhu J, Liu B, Xie Q, Yuan T, Wu Y, Chen R, Xie X et al. Apigenin suppresses mycoplasma-induced alveolar macrophages necroptosis via enhancing the methylation of TNF-α promoter by PPARγ-Uhrf1 axis. Phytomedicine. 2023;108:154504. https://doi.org/10.1016/j.phymed.2022.154504.
44. Kanwal R, Datt M, Liu X, Gupta S. Dietary Flavones as Dual Inhibitors of DNA Methyltransferases and Histone Methyltransferases. PLoS One. 2016;11:e0162956. https://doi.org/10.1371/journal.pone.0162956.
45. Bonilla-Vidal L, Świtalska M, Espina M, Wietrzyk J, García ML, Souto EB, Gliszczyńska A, Sánchez López E. Dually Active Apigenin-Loaded Nanostructured Lipid Carriers for Cancer Treatment. Int J Nanomedicine. 2023;18:6979-97. https://doi.org/10.2147/ijn.S429565.
46. Borges G, Fong RY, Ensunsa JL, Kimball J, Medici V, Ottaviani JI, Crozier A. Absorption, distribution, metabolism and excretion of apigenin and its glycosides in healthy male adults. Free Radic Biol Med. 2022;185:90-96. https://doi.org/10.1016/j.freeradbiomed.2022.04.007.

Figure Legend

Figure 1 Flowchart of the study.

Figure 2 Construction of WGCNA Notes.

(A) Hierarchical clustering of the GSE48213 samples to identify potential outliers. The top 5,000 genes ranked by standard deviation were included in the weighted gene co-expression network analysis. (B) Scale-free topology fit index across different soft-thresholding powers. (C) Mean connectivity across different soft-thresholding powers. (D) Gene clustering dendrogram with module assignments indicated by different colors. (E) Heatmap of the topological overlap matrix. Different colors on the horizontal and vertical axes represent different modules, and increasing color intensity indicates stronger connectivity between genes.

Figure 3 Identification of key modules.

(A) Module–trait relationship heatmap showing the correlations between module eigengenes and BC molecular subtypes. Each row represents a module eigengene, each column represents a molecular subtype, and each cell displays the corresponding correlation coefficient and P value. (B–D) Scatter plots of module membership versus gene significance, together with the corresponding gene-expression heatmaps, for the turquoise, yellow, and brown modules, respectively.

Figure 4 Identification of key genes.

DEGs of apigenin before and after treatment of breast cancer (GSE119552). Heatmap (A) and a volcano plot (B). (C) Identification of common genes between DEGs and the key modules (turquoise, yellow, and brown modules) by overlapping them. The 101 key genes in the three modules were also DEGs in GSE119552. (D-E) Functional annotation of key genes: Gene Ontology and pathway enrichment analysis of key genes by DAVID and visualisation by R. (F) PPI network analysis of key genes. (G) Ten hub genes were obtained from CytoHubba with PPI network analysis.

Figure 5 Validation of hub genes at the transcriptional level.

(A-J) Validation of hub genes using TCGA-BRCA data in GEPIA showed that the expression status of 10 hub genes was correlated with disease status, indicating they play an essential role in promoting tumorigenesis.

Figure 6 IHC of the 9 hub genes based on the Human Protein Atlas.

(A-I) Validation of hub genes by the Human Protein Atlas database. There were no related IHC samples of EXOSC5 in the database. The translational expression levels of the remaining 9 hub genes were correlated with disease status, as they are differentially expressed in BC samples compared with non-cancerous samples.

Figure 7 Survival analysis of hub genes.

(A-J) Kaplan-Meier curves of key genes in patients with BC. The red and blue curves represent the High-expression and Low- expression groups, respectively. 10 hub genes with the highest node degrees and significant results from the survival

analysis. They were MRTO4, BYSL, RRS1, NOP2, RRP9, MYBBP1A, EXOSC5, METTL1, NOP16, IPO4, respectively.

Figure 8 Identification of hub genes. ROC curve analysis of the candidate hub genes in the GSE33447 dataset.

Figure 9 Association of 10 hub gene expression with tumor purity and immune cells using the TCGA-BRCA dataset.

Figure 10 Methylation analyses of hub genes.

(A-J) Methylation levels of MRTO4, BYSL, RRS1, NOP2, RRP9, MYBBP1A, EXOSC5, METTL1, NOP16 and IPO4 in breast cancer and normal tissues.

Figure 11 Genetic alterations associated with hub genes in TCGA-BRCA.

(A) Summary of the frequencies and types of genetic alterations in the 10 hub genes in patients from the TCGA-BRCA cohort. (B) Frequencies and types of genetic alterations across BC subtypes. (C–J) Detailed mutation profiles of MRTO4, BYSL, RRS1, NOP2, RRP9, MYBBP1A, EXOSC5, and IPO4, respectively.

Figure 12 Molecular docking of the interactions between apigenin and 10 target proteins.

Figure 13. Apigenin inhibits the viability and downregulates the expression of IPO4, METTL1 and RRP9 in BC cells in *vitro*

(A–C) The viability of MCF-7, MDA-MB-231, and SK-BR-3 cells after 24 h treatment with Apigenin by CCK-8 assay. (D–F) Calculated IC_{50} values for apigenin in three BC cell lines.

(G-I) Representative Western blot images of IPO4, METTL1, RRP9, and internal reference GAPDH in MCF-7, MDA-MB-231, and SK-BR-3 cells treated with apigenin. (J-L) Quantitative grayscale analysis for relative protein expression levels. All quantitative data are presented as the mean ± SD from three independent biological replicates (n = 3). Statistical differences among multiple groups were evaluated using one-way analysis of variance followed by Tukey's multiple-comparisons test. ns, not significant; $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$ versus the control group.

Figure 14. Rescue experiments verify the regulatory effect of apigenin on IPO4, METTL1, and RRP9.

(A) Schematic diagram of the rescue experiment. (B-D) Representative Western blot bands of IPO4, METTL1, RRP9, and the internal reference GAPDH in MCF-7, MDA-MB-231, and SK-BR-3 cells, respectively. (E-G) Quantitative grayscale analysis of the relative protein expression levels of target bands. Quantitative data are presented as the mean ± SD from three independent biological replicates (n = 3). Statistical differences among multiple groups were evaluated using one-way analysis of variance followed by Tukey's multiple-comparisons test. ns, not significant; $*P < 0.05$ versus the control group.

Figure 1

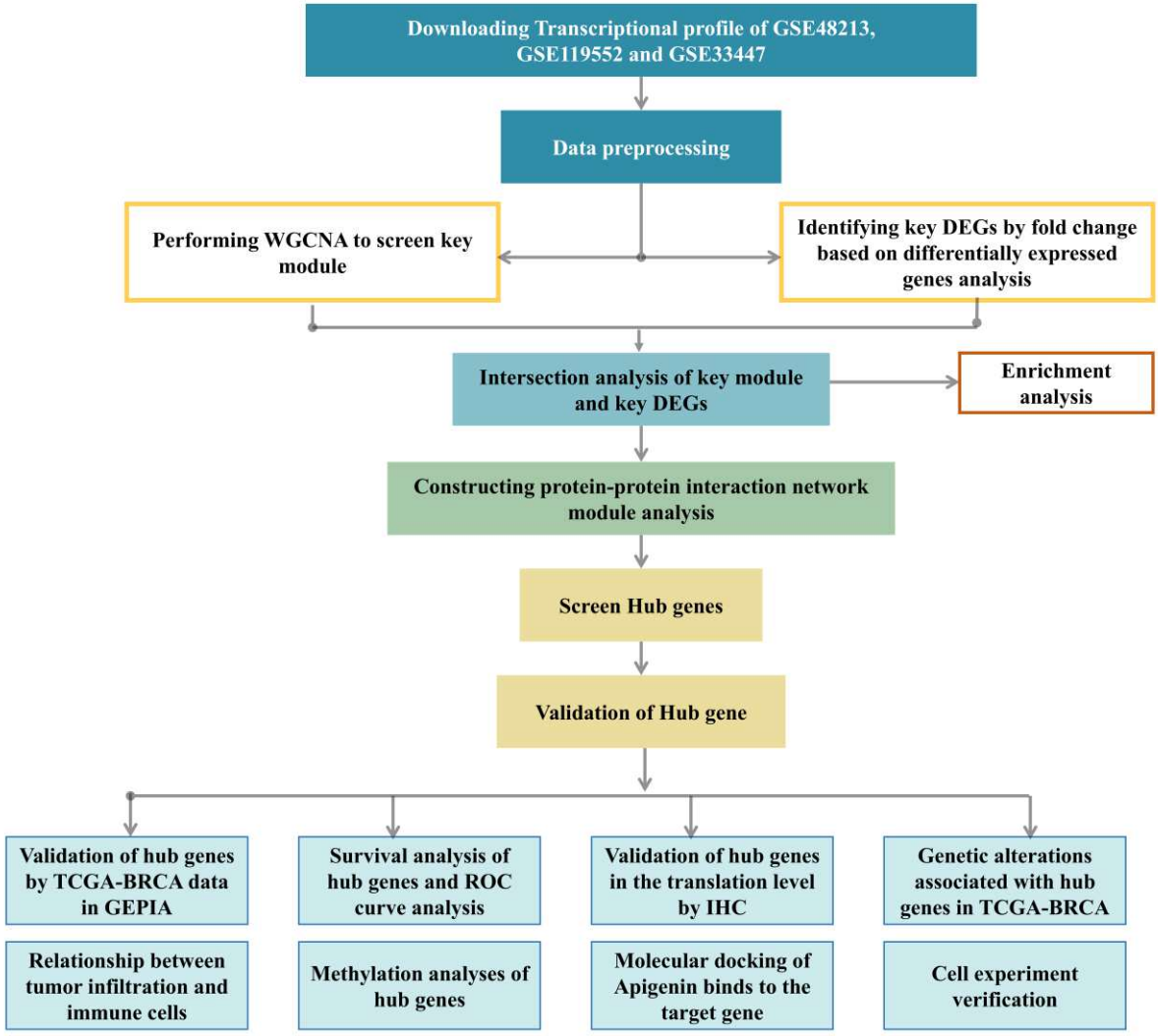


Figure 2

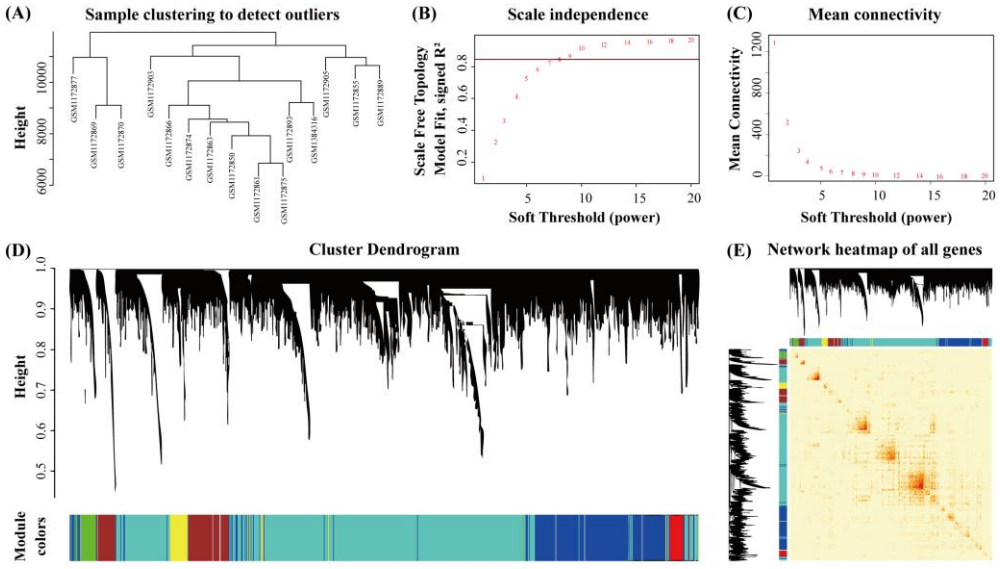


Figure 3

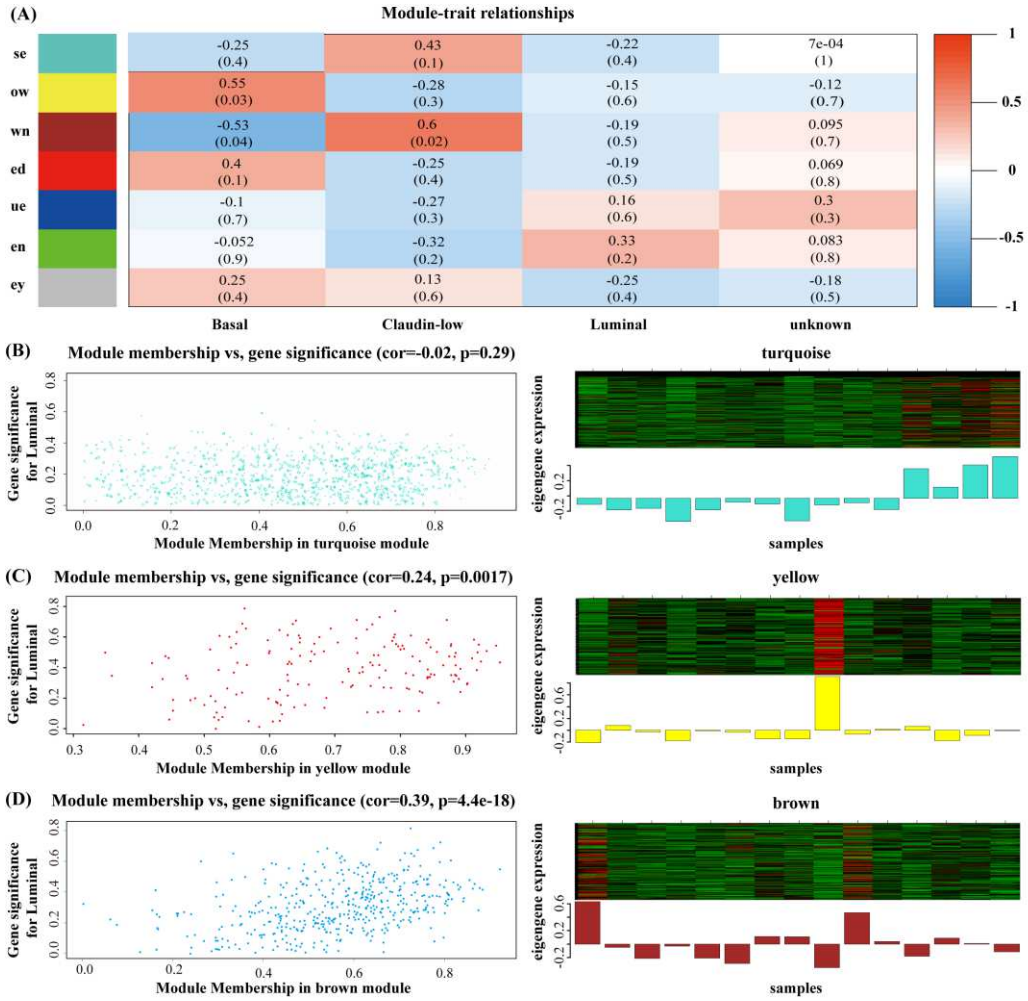

(A)
Module-trait relationships
se
ow
wn
ed
ue
en
ey
-0.25 (0.4)
0.43 (0.1)
-0.22 (0.4)
7e-04 (1)
0.55 (0.03)
-0.28 (0.3)
-0.15 (0.6)
-0.12 (0.7)
-0.53 (0.04)
0.6 (0.02)
-0.19 (0.5)
0.095 (0.7)
0.4 (0.1)
-0.25 (0.4)
-0.19 (0.5)
0.069 (0.8)
-0.1 (0.7)
-0.27 (0.3)
0.16 (0.6)
0.3 (0.3)
-0.052 (0.9)
-0.32 (0.2)
0.33 (0.2)
0.083 (0.8)
0.25 (0.4)
0.13 (0.6)
-0.25 (0.4)
-0.18 (0.5)
Basal
Claudin-low
Luminal
unknown
1
0.5
0
-0.5
-1
(B)
Module membership vs, gene significance (cor=-0.02, p=0.29)
Gene significance for Luminal
Module Membership in turquoise module
turquoise
eigengene expression
samples
(C)
Module membership vs, gene significance (cor=0.24, p=0.0017)
Gene significance for Luminal
Module Membership in yellow module
yellow
eigengene expression
samples
(D)
Module membership vs, gene significance (cor=0.39, p=4.4e-18)
Gene significance for Luminal
Module Membership in brown module
brown
eigengene expression
samples

Figure 4

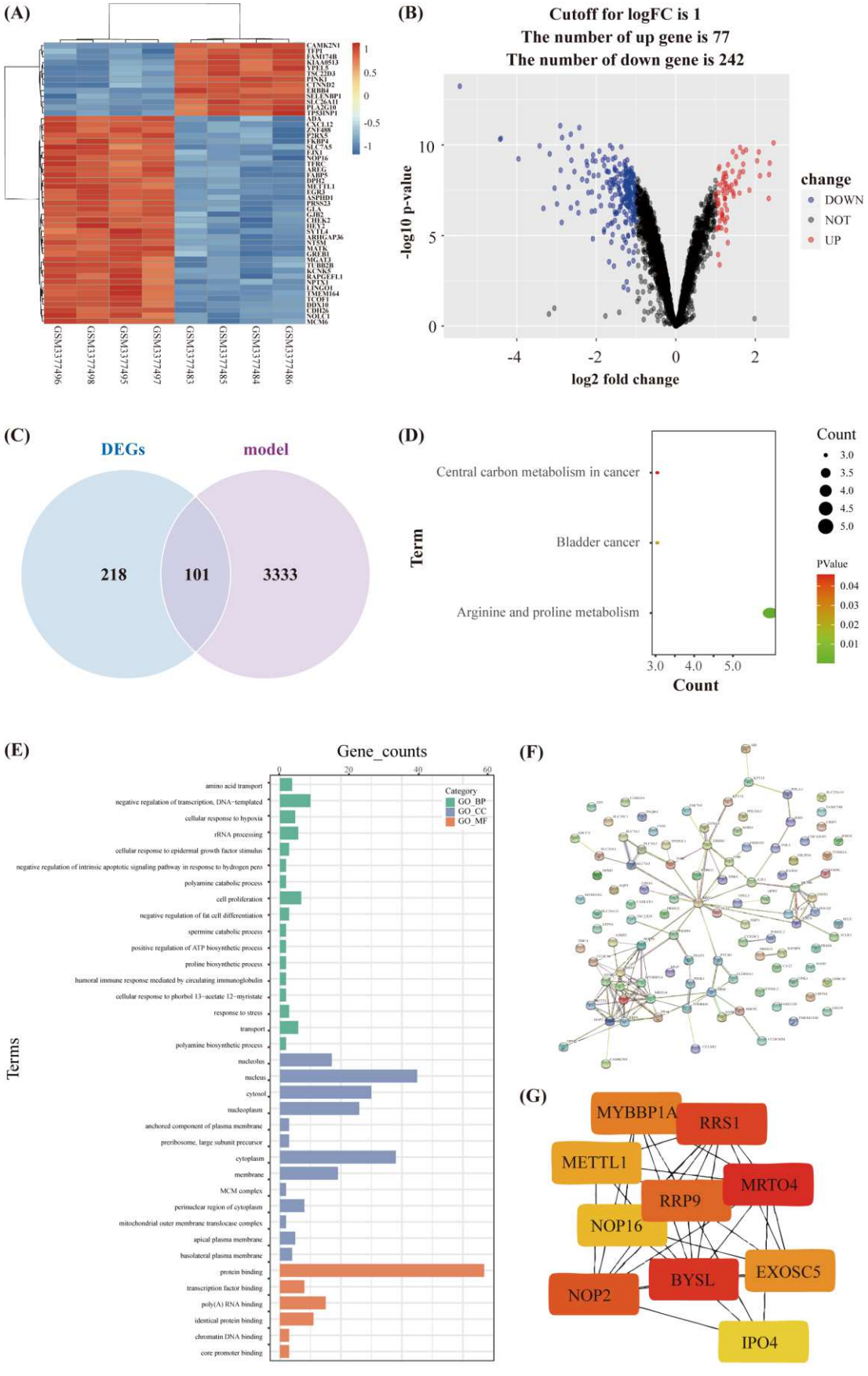
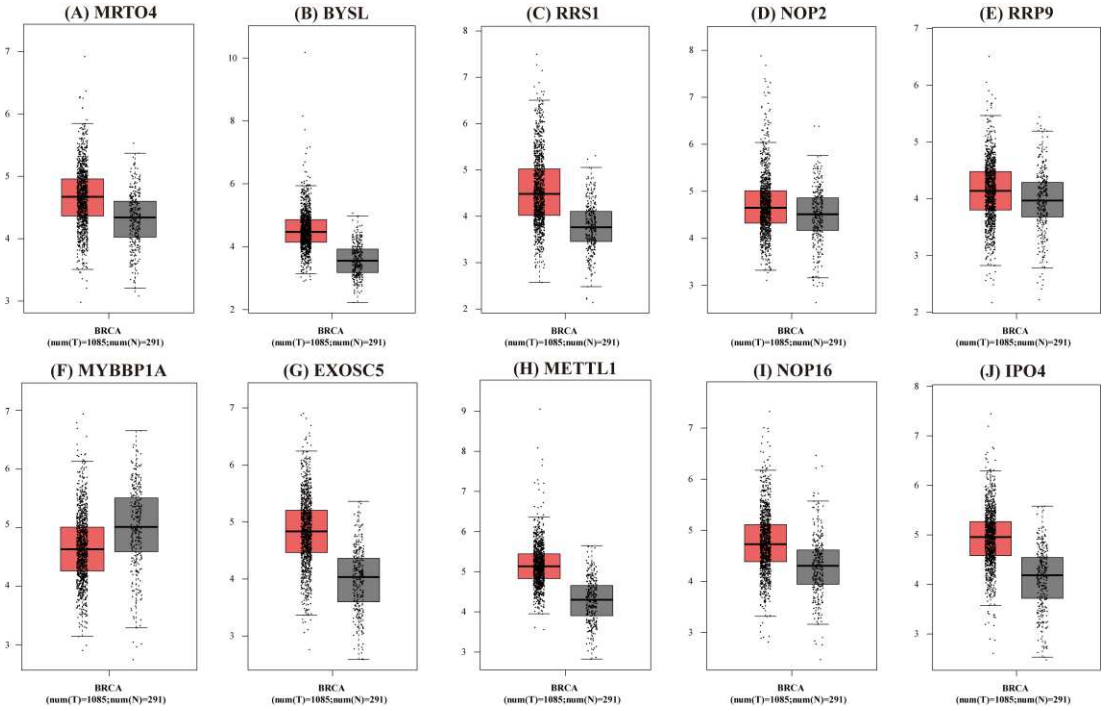

(A)
GSM3377496
GSM3377498
GSM3377495
GSM3377497
GSM3377483
GSM3377485
GSM3377484
GSM3377486
(B)
Cutoff for logFC is 1
The number of up gene is 77
The number of down gene is 242
-log10 p-value
log2 fold change
change
DOWN
NOT
UP
(C)
DEGs
model
218
101
3333
(D)
Central carbon metabolism in cancer
Bladder cancer
Arginine and proline metabolism
Term
Count
PValue
(E)
Gene_counts
Terms
Category
GO_BP
GO_CC
GO_MF
amino acid transport
negative regulation of transcription, DNA-templated
cellular response to hypoxia
rRNA processing
cellular response to epidermal growth factor stimulus
polyamine catabolic process
cell proliferation
negative regulation of fat cell differentiation
spermine catabolic process
positive regulation of ATP biosynthetic process
proline biosynthetic process
humoral immune response mediated by circulating immunoglobulin
cellular response to phorbol 13-acetate 12-myristate
response to stress
transport
polyamine biosynthetic process
nucleolus
nucleus
cytosol
nucleoplasm
anchored component of plasma membrane
preribosome, large subunit precursor
cytoplasm
membrane
MCM complex
perinuclear region of cytoplasm
mitochondrial outer membrane translocase complex
apical plasma membrane
basolateral plasma membrane
protein binding
transcription factor binding
poly(A) RNA binding
identical protein binding
chromatin DNA binding
core promoter binding
(F)
(G)
MYBBP1A
RRS1
METTL1
MRTO4
RRP9
NOP16
EXOSC5
BYSL
NOP2
IPO4

Figure 5



Figue 6

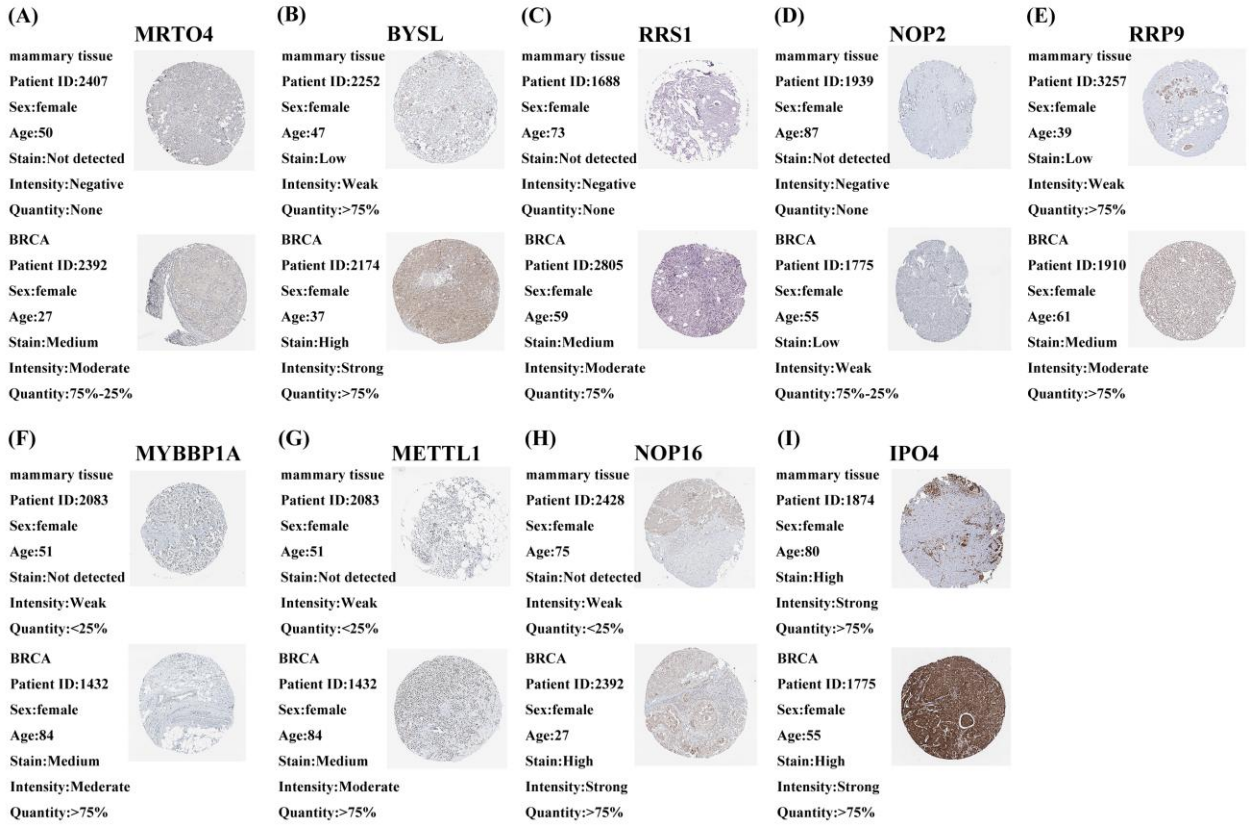


Figure 7

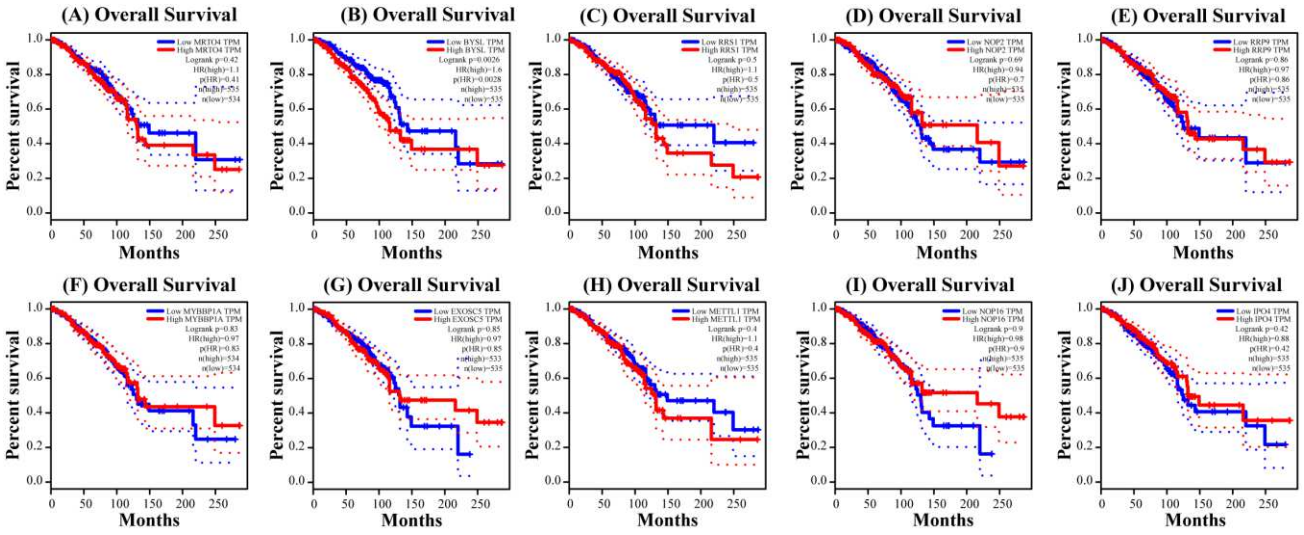


Figure 8

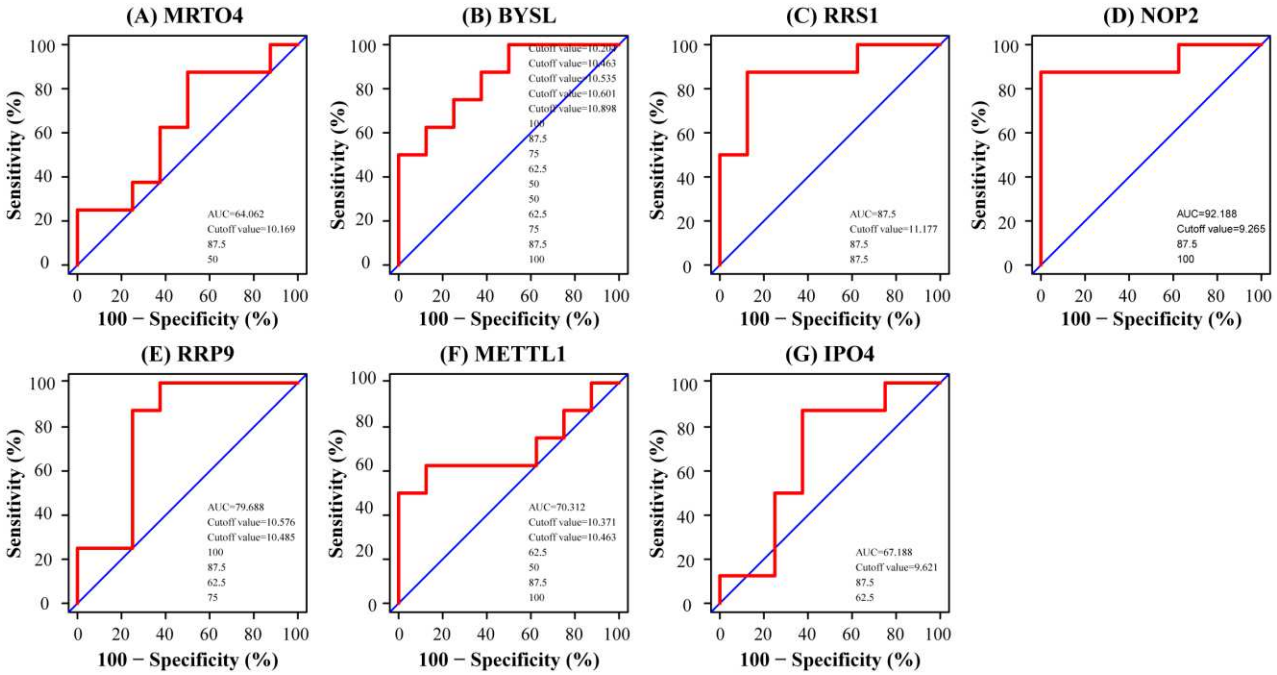


Figure 9

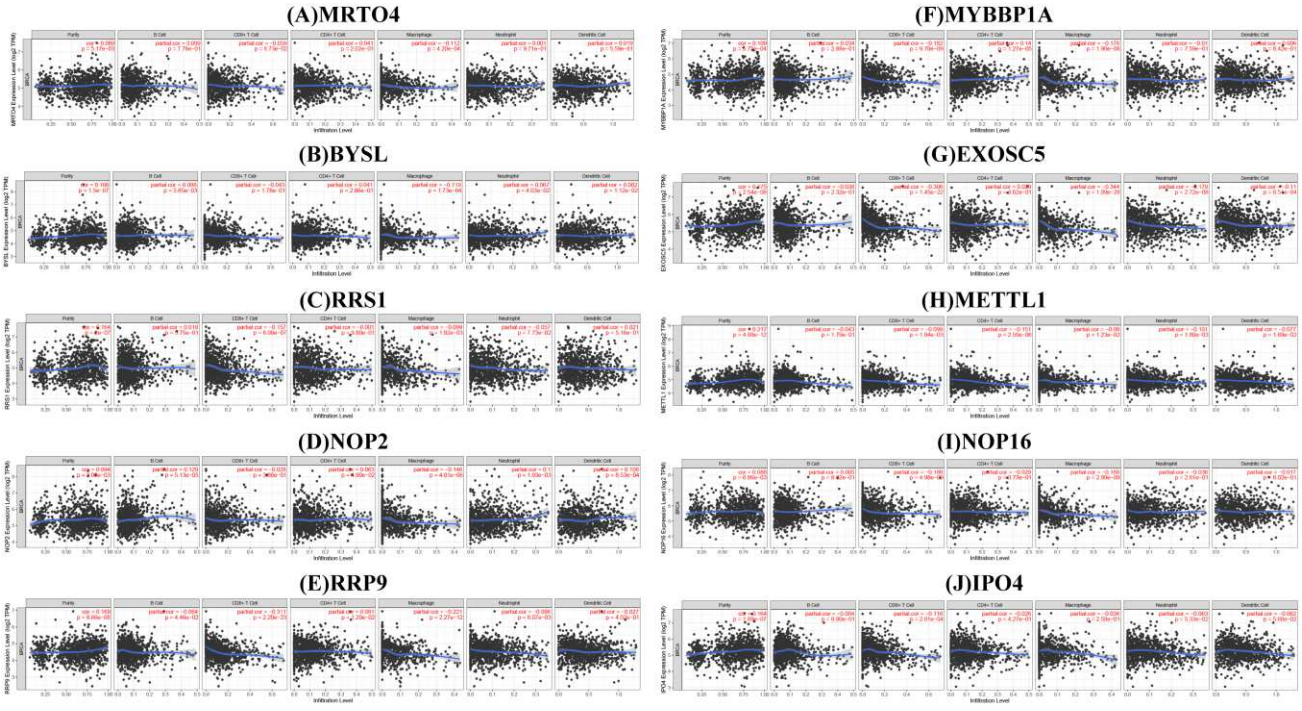


Figure 10

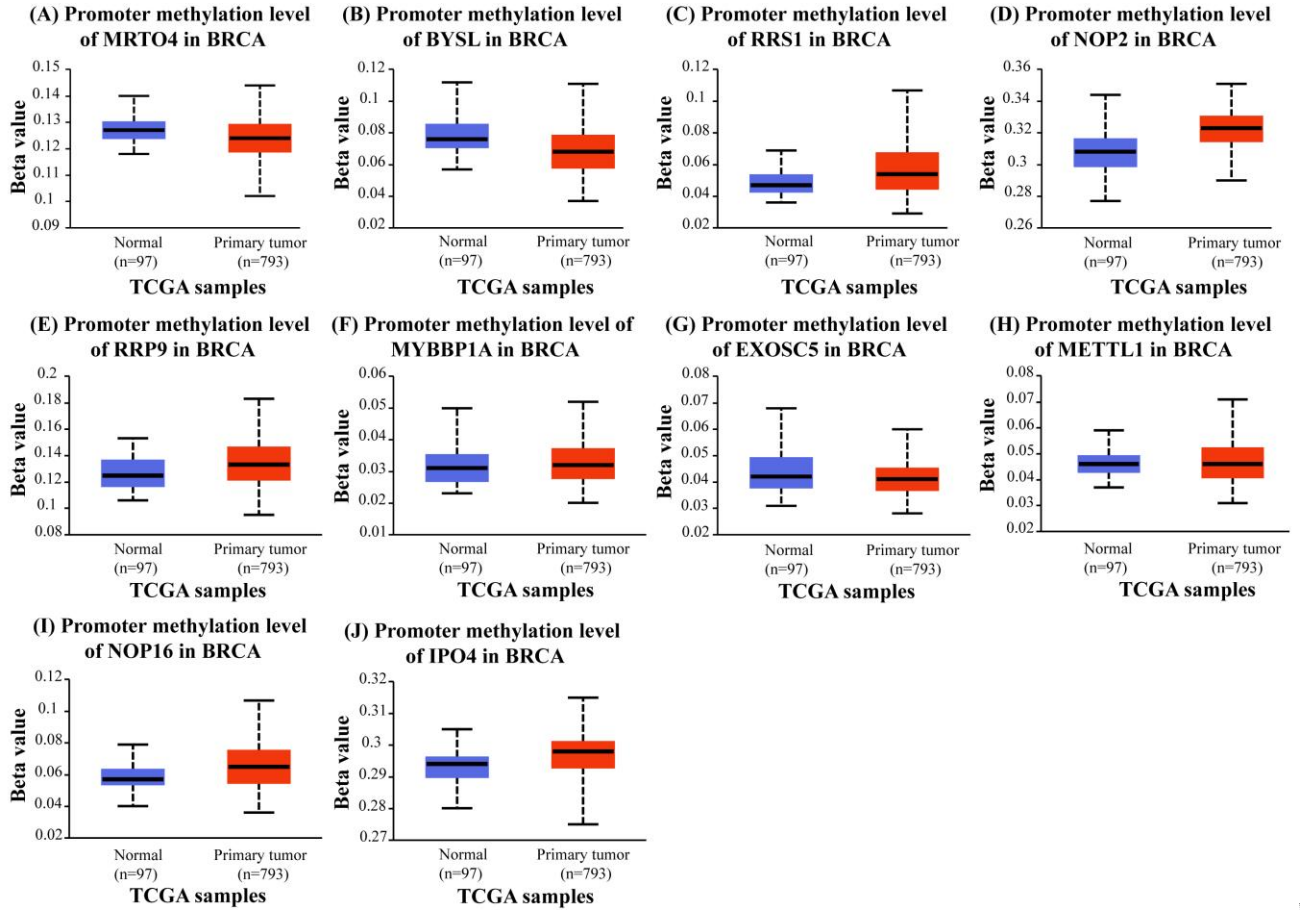


Figure 11

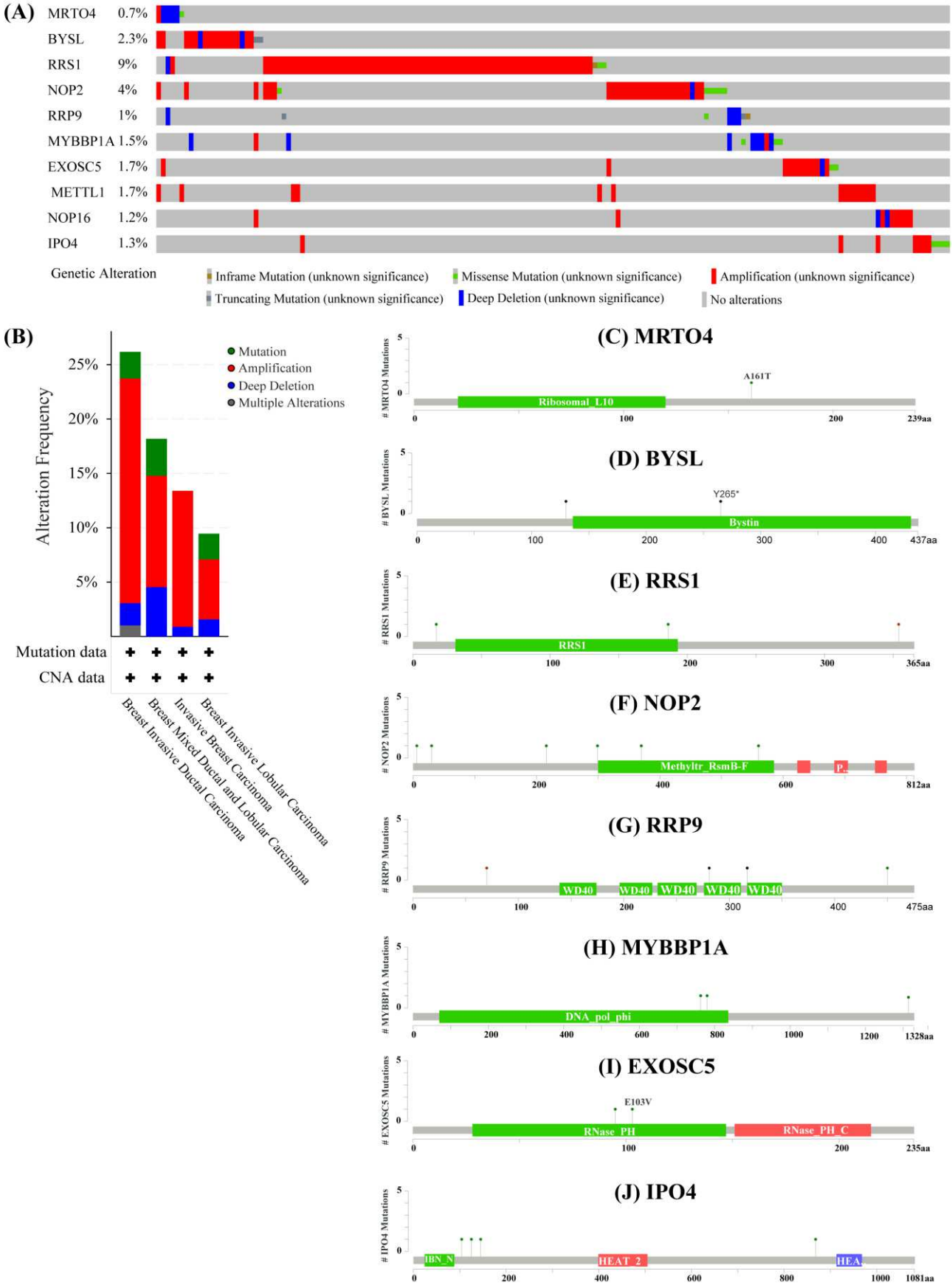

(A)
MRTO4 0.7%
BYSL 2.3%
RRS1 9%
NOP2 4%
RRP9 1%
MYBBP1A 1.5%
EXOSC5 1.7%
METTL1 1.7%
NOP16 1.2%
IPO4 1.3%
Genetic Alteration
Inframe Mutation (unknown significance)
Missense Mutation (unknown significance)
Amplification (unknown significance)
Truncating Mutation (unknown significance)
Deep Deletion (unknown significance)
No alterations
(B)
Alteration Frequency
25%
20%
15%
10%
5%
Mutation
Amplification
Deep Deletion
Multiple Alterations
Mutation data
CNA data
Breast Invasive Ductal Carcinoma
Breast Mixed Ductal and Lobular Carcinoma
Invasive Breast Carcinoma
Breast Invasive Lobular Carcinoma
(C) MRTO4
MRTO4 Mutations
A161T
Ribosomal_L10
239aa
(D) BYSL
BYSL Mutations
Y265*
Bystin
437aa
(E) RRS1
RRS1 Mutations
RRS1
365aa
(F) NOP2
NOP2 Mutations
Methyltr_RsmB-F
812aa
(G) RRP9
RRP9 Mutations
WD40
475aa
(H) MYBBP1A
MYBBP1A Mutations
DNA_pol_phi
1328aa
(I) EXOSC5
EXOSC5 Mutations
E103V
RNase_PH
RNase_PH_C
235aa
(J) IPO4
IPO4 Mutations
IBN_N
HEAT_2
1081aa

Figure 12

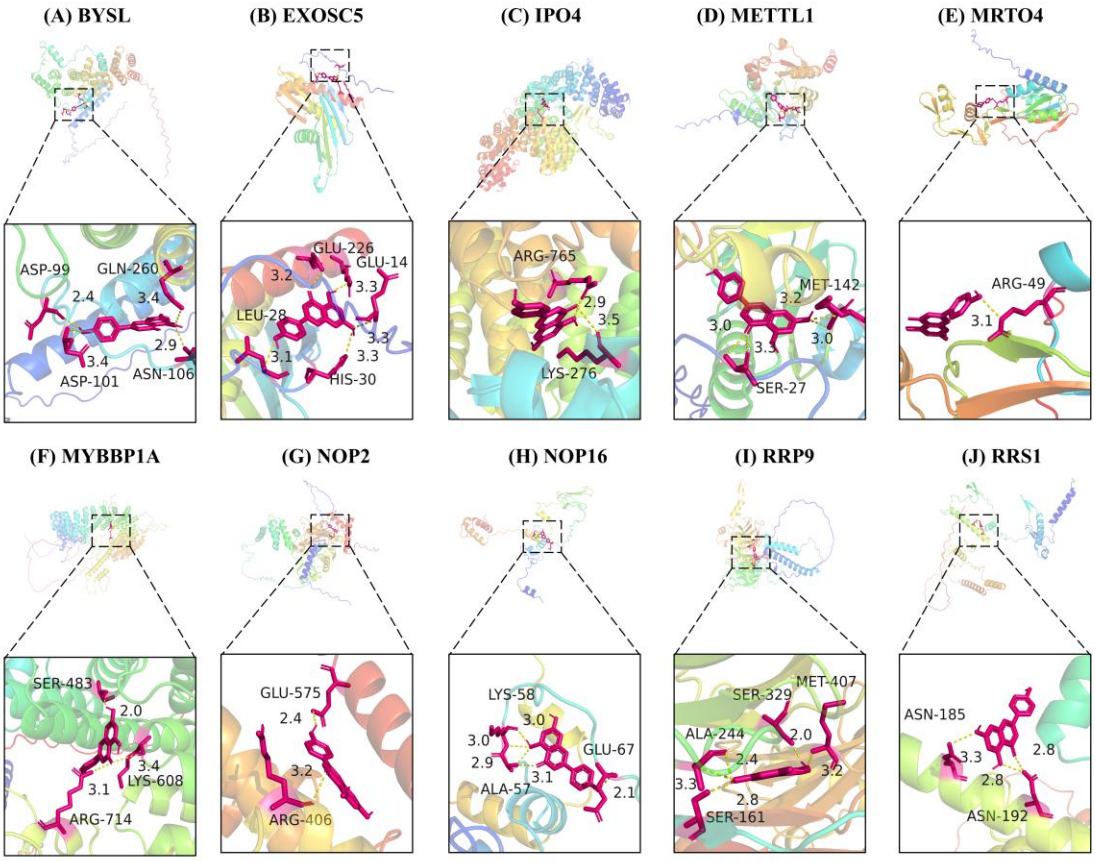

(A) BYSL
ASP-99
GLN-260
2.4
3.4
2.9
3.4
ASP-101
ASN-106
(B) EXOSC5
GLU-226
GLU-14
3.2
3.3
LEU-28
3.3
3.1
3.3
HIS-30
(C) IPO4
ARG-765
2.9
3.5
LYS-276
(D) METTL1
MET-142
3.2
3.0
3.3
3.0
SER-27
(E) MRTO4
ARG-49
3.1
(F) MYBBP1A
SER-483
2.0
3.4
LYS-608
3.1
ARG-714
(G) NOP2
GLU-575
2.4
3.2
ARG-406
(H) NOP16
LYS-58
3.0
3.0
GLU-67
2.9
3.1
ALA-57
2.1
(I) RRP9
SER-329
MET-407
ALA-244
2.0
2.4
3.2
3.3
2.8
SER-161
(J) RRS1
ASN-185
3.3
2.8
2.8
ASN-192

Figure 13

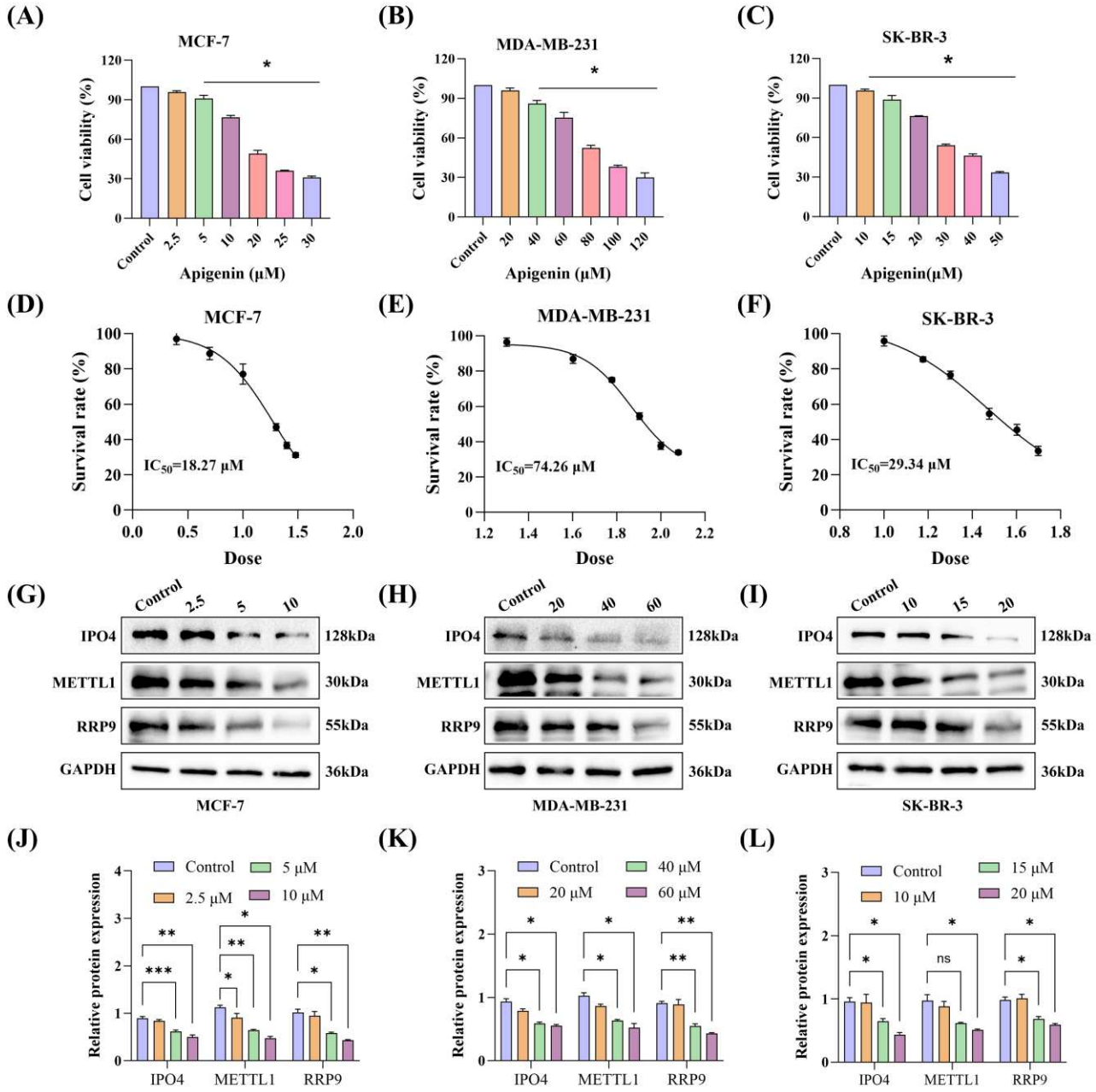

(A)
MCF-7
Cell viability (%)
Apigenin (μM)
(B)
MDA-MB-231
Cell viability (%)
Apigenin (μM)
(C)
SK-BR-3
Cell viability (%)
Apigenin(μM)
(D)
MCF-7
Survival rate (%)
IC50=18.27 μM
Dose
(E)
MDA-MB-231
Survival rate (%)
IC50=74.26 μM
Dose
(F)
SK-BR-3
Survival rate (%)
IC50=29.34 μM
Dose
(G)
IPO4
METTL1
RRP9
GAPDH
128kDa
30kDa
55kDa
36kDa
MCF-7
(H)
MDA-MB-231
(I)
SK-BR-3
(J)
Relative protein expression
Control
2.5 μM
5 μM
10 μM
IPO4
METTL1
RRP9
(K)
20 μM
40 μM
60 μM
(L)
10 μM
15 μM
20 μM
ns

Figure 14

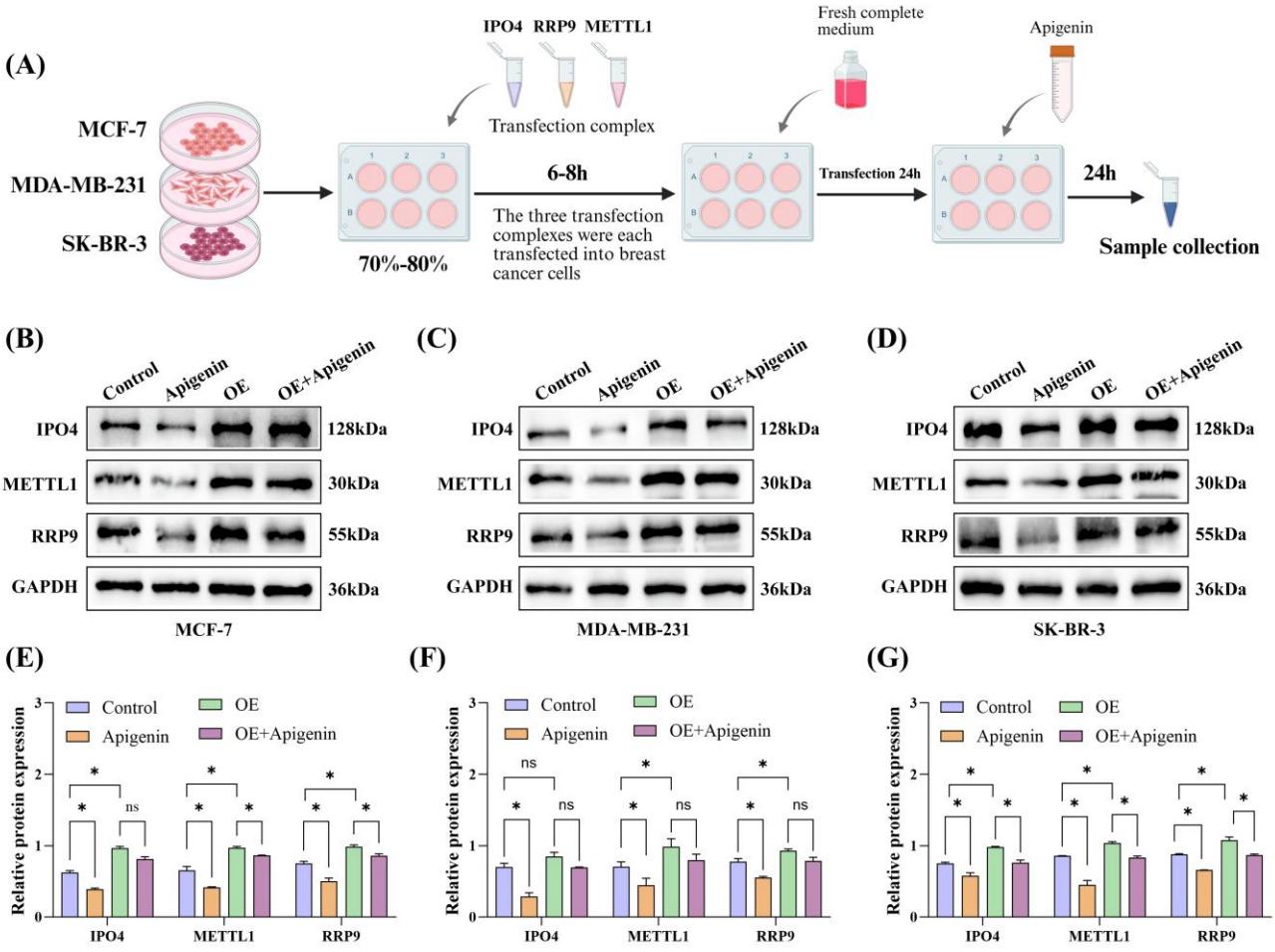

(A)
MCF-7
MDA-MB-231
SK-BR-3
70%-80%
IPO4 RRP9 METTL1
Transfection complex
6-8h
The three transfection complexes were each transfected into breast cancer cells
Fresh complete medium
Transfection 24h
Apigenin
24h
Sample collection
(B)
(C)
(D)
Control
Apigenin
OE
OE+Apigenin
IPO4
METTL1
RRP9
GAPDH
128kDa
30kDa
55kDa
36kDa
MCF-7
MDA-MB-231
SK-BR-3
(E)
(F)
(G)
Relative protein expression
IPO4
METTL1
RRP9
ns

Table 1 The binding energy and hydrogen-bonded amino acids of apigenin after docking with 10 target gene molecules

Target gene	UniProt KB	binding energy （kcal/mol）	Hydrogen-bonded amino acids
BYSL	Q13895	-7.6	ASP-99, GLN-260, ASP-101, ASN-106
EXOSC5	Q9NQT4	-7.8	GLU-226, LEU-28, GLU-14, HIS-30
IPO4	Q8TEX9	-8.0	ARG-765, LYS-276
METTL1	Q9UBP6	-8.6	MET-142, SER-27
MYBBP1A	Q9BQG0	-7.3	SER-483, LYS-6088, ARG-714
MRTO4	Q9UKD2	-6.2	ARG-49
NOP2	P46087	-7.8	GLU-575, ARG-406
RRS1	Q15050	-6.5	ASN-185, ASN-192
RRP9	O43818	-8.8	SER-329, ALA-244, MET-407, SER-161
NOP16	Q9Y3C1	-7.1	LYS-58, GLU-67, ALA-57

Supplementary Files

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

- Fig13Westernblot.docx
- Fig14Westernblot.docx