

Systematic identification of drug-based prognostic biomarkers by integrating multiomics in breast cancer

Hongying Zhao

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China

Lei Yu

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China

Yanyan Ping

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China.

Shihua Lin

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China

Caiyu Zhang

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China

Haotian Xu

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China

Tengyue Li

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China

Waidong Huang

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China

Junjie Lei

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China

Zhijun Leng

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China

Jing Li

Department of Ultrasonic medicine, The 1st Affiliated Hospital of Heilongjiang University of Chinese Medicine, Harbin 150040, China

Shangwei Ning

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China.

Li Wang (✉ wangli@hrbmu.edu.cn)

Harbin Medical University

Research

Keywords: prognostic biomarkers, drug repositioning, breast cancer, multi-omics, human function linkages network

Posted Date: May 11th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-26796/v1>

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

[Read Full License](#)

Abstract

Background

Drug repositioning plays an important role in the current drug research and development field, as the high cost and long period required to bring drugs into commerce are driving efforts to repurpose drugs approved by the U.S. Food and Drug Administration for new uses.

Methods

Here, we used breast cancer as a disease model and systematically proposed a new strategy based on multidimensional omics data (e.g., mRNA expression, DNA methylation, DNA copy number, and mutational profiles) to reconstruct the human functional linkage network. In addition, it was used to systematically interpret the functional relationship between human genes and to establish a measure of the relationship between drugs and diseases to reset drugs and achieve drug-based prognostic biomarkers.

Results

We found that differentially expressed genes driven by DNA methylation, DNA copy number, and somatic mutation were significantly enriched in cancer-related biological pathways and functions. Survival analysis showed that multiomics-driven target genes were repositioned, which significantly distinguished between high-risk and low-risk groups in breast cancer patients, such as the methylation-driven gene glycoprotein M6B ($P = 0.00141$), copy number-driven gene acyl-CoA synthetase long chain family member 1 ($P = 0.0582$), mutation-driven gene minichromosome maintenance 10 replication initiation factor ($P = 0.0303$), and multiomics synergy driven gene catechol-O-methyltransferase domain containing 1 ($P = 0.0071$).

Conclusions

Our approach provides a basis for identifying novel candidate drug and drug-based prognostic biomarkers.

Background

The high cost and long period required[1] to bring drugs into commerce are driving efforts to repurpose U.S Food and Drug Administration (FDA)-approved drugs to find new uses, thereby reducing the overall cost of commercialization and shortening the gap between drug discovery and availability[2]. The most successful example of drug repositioning is sildenafil, which was originally developed as a cardiovascular drug and was later repurposed to treat erectile dysfunction[3]. Similarly, zidovudine was

originally developed as an anticancer drug and was later repositioned to treat human immunodeficiency virus[4]. The discovery of new uses for these drugs, although serendipitous, has motivated research into more systematic approaches that may markedly amplify the number of discoveries.

Systematic approaches generally begin with some form of computer-based screening to generate large numbers of plausible candidates. Many current computational strategies are exploited for drug repositioning[5–7]. Drug similarities include drug-induced phenotypic side effects, chemical structure[8], and molecular activity[9, 10]. Disease similarities include semantic phenotypic similarity and phenotypic similarity constructed using similarities in medical subject headings in the Online Mendelian Inheritance in Man database[11, 12]. Because disease biomarkers have not changed, the efficacy of drug candidates produced by these methods will not exceed existing drugs. A more general approach is to determine the relationship between disease- and drug-induced transcriptional profiles, and to analyze the interaction between FDA-approved drugs and disease gene expression profiles, thereby determining if the drug is a candidate for repositioning[13]. Because this method only uses non-functionally related differentially expressed genes (DEGs) that interact with each other, it is not able to well interpret the functional linkages between interfering genes, leading to limitations in the identified drug candidates.

In this study, we used breast cancer as a disease model to systematically propose a new strategy for drug repositioning and to identify drug-associated biomarkers based on reconstruction of the human functional linkage network (FLN). The goal of the study was to find disease risk genes that have strong functional relationships with breast cancer drug target genes, in order to reconstruct human FLNs for drug relocation. In particular, we searched for drugs that strongly perturb genes with the following properties: strong functional connections; abnormal changes in the molecule during the disease stage (e.g., mutation, DNA copy number variation, DNA methylation); expression levels that are strongly interfered with by the disease; and significant enrichment in biological pathways affected by disease. Our system interpreted functional relationships among human genes and established a measure of the relationship between drugs and diseases to identify drug-based prognostic biomarkers.

Data And Methods

A total of 304 breast cancer disease samples and 18 normal samples and their clinical information (e.g., age, survival, gender, and time) based on The Cancer Genome Atlas (TCGA; <http://cancergenome.nih.gov/>) database as well as mRNA, DNA methylation levels, copy number variation, and somatic mutations were simultaneously detected in each sample. Based on the Drug Bank (<https://www.drugbank.ca/>) database, we obtained FDA-approved drugs and their target genes (breast cancer drug response gene [DRG]), which were clearly described for the treatment of breast cancer. A human FLN was obtained from the visANT (<http://visant.bu.edu/misi/fln/>) database, which provides a quantitative measure of the degree of functional association between any pair of human genes[14].

Identification Of Breast Cancer Risk Genes Based On Multiomics

The gene expression profile detected by Agilent 244K contained 17,815 genes, and the DNA methylation data detected by the Infinium HumanMethylation27 BeadChip Kit contained methylation levels of 27,578 CpG sites associated with 14,475 gene promoters. The methylation level of each CpG site was calculated as the ratio of methylated probe signal to the sum of methylated and unmethylated probes ranging from 0 (unmethylated) to 1 (methylated) continuous value. The copy number spectrum was derived from the SNP Array 6.0 chip, and the copy number region was segmented by the cyclic binary segmentation method; log₂ was used for numerical conversion [15]. We obtained somatic mutation data from TCGA and obtained a total of 8,543 somatic mutation genes by removing silent mutations. For mRNA expression profiles and DNA methylation profiles, we used fold change and *t*-test (false discovery rate [FDR] < 0.001) to identify DEGs with significantly different mRNA changes and methylation abnormalities. Because of the large difference between the disease sample and the normal sample size, we performed 1,000 random samplings and selected genes with a frequency greater than 10% as DEGs and DEGs with abnormal methylation changes. For DNA copy number data, we used the Genomic Identification of Significant Targets in Cancer (GISTIC; version 2) algorithm to identify genes with genomic amplification and deletion in breast cancer with default parameters. Finally, we selected DEGs with DNA methylation, copy number, and somatic mutation and consistent mRNA expression as risk genes for breast cancer.

Identifying breast cancer risk genes with significantly enriched biological pathways and functions

We performed pathway enrichment analysis of breast cancer risk genes based on the KOBAS 3.0 web server for gene/protein functional annotation and selected pathways with an FDR < 0.05 as significantly enriched biological pathways [16]. Similarly, we used DAVID to perform functional enrichment analysis of breast cancer risk genes, and selected functions with an FDR < 0.05 as significantly enriched biological functions [17].

Reconstructing Human Functional Linkage Network

We mapped the significant differences in DNA methylation, copy number variation, mutation-affected breast cancer risk genes, and DRG into human FLNs. Accordingly, we obtained and integrated three reconstructed human FLNs containing information on the multiomics of diseases.

Identifying Reset Drug Targets Based On Drugs And Disease

We analyzed each reconstructed human FLN as follows. For each mutation-driven DEG (e.g., gene A), we identified one-step neighbors to which they were linked (only DEGs and DRGs were considered); if the DEG was the same gene as the DRG, it was removed from the network and assigned a DD score of 1. DRGs were sorted according to the weight of the reconstructed human FLN and then their AUC_A values

(threshold t ; DRG above the t value was true positive, and DRG below the t value was true negative) were calculated. For each DRG (e.g., gene B), we took the same approach to obtain its AUC_B value (the mutation-driving gene above t value was true positive, and the DRG below the t value was true negative). Finally, the correlation between each DEG and the DRG was defined as the geometric mean of the DEG and DRG AUC values [18]:

$$DD \text{ Score } AB = \sqrt{AUC_A \times AUC_B}$$

We sorted the DD scores and selected the top 100 gene pairs. Through literature mining, we considered the drug corresponding to each breast cancer risk gene as the reset drug target (Fig. 1).

Prognosis Analysis Using Reset Drug Targets In Breast Cancer

The reset drug target for breast cancer was determined using the DD Score method, and the survival and survminer R packages were used to perform the survival analysis. To assess the prognostic efficacy of breast cancer reset drug targets, we took the median value of all of the disease risk genes and divided them into two groups: the high-expression group had expression values greater than or equal to the median, the low-expression group had expression values less than the median. Survival analysis was performed by dividing the two sets of gene expression values and combining with clinical data.

Results

Identifying DEGs in breast cancer based on multiomics data

For mRNA, DNA methylation, copy number variation, and somatic mutation, we identified 2,388, 1,442, 3,348, and 8,543 DEGs, respectively. Among them, there were 1,020 highly expressed genes, 1,368 low-expressed genes, 670 significantly hypermethylated genes, 770 significantly hypomethylated genes, 715 copies of copy-numbered genes, and 2,633 copies of deleted genes. The numbers of upregulated DEGs with DNA methylation, DNA copy number, and somatic mutations consistent with mRNA expression were 35, 30, and 138, respectively; the number of downregulated DEGs was 76, 89, and 359, respectively (Fig. 2).

Identifying significantly enriched biological pathways and functions based on DEGs

We performed pathway enrichment analysis of methylation, copy number, and mutant disease risk genes, and took the significantly enriched pathway TOP20 as an example (Fig. 3A). The results showed that DNA methylation, copy number, and mutation-driven breast cancer risk genes were significantly enriched in “pathways in cancer,” “microRNAs (miRNAs) in cancer,” “PI3K-Akt signaling pathway,” and “cell

adhesion molecules (CAMs).” PI3K-Akt is a signaling pathway for a variety of cancers that regulates many basic cellular functions such as DNA transcription, protein translation, cell proliferation, cell growth, and survival [19]. It also controls key cellular processes of cancer through phosphorylation. CAM is a glycoprotein that plays a key role in a variety of biological processes [20, 21]. Existing studies have shown that expression of CAM is associated with cancer [22–24]. We performed functional enrichment analysis on methylation, copy number, and mutant disease risk genes. The results showed that DNA methylation, copy number, and mutation-driven breast cancer risk genes were significantly enriched in the “negative regulation of apoptotic process” and “positive regulation of cell proliferation/negative regulation of cell proliferation.” Apoptosis is widely recognized as a positive process for the prevention and treatment of cancer, but apoptosis may also lead to adverse reactions and may even promote cancer. It has been found that many non-coding RNAs regulate apoptosis and play a positive role in tumor cell therapy. Cell proliferation is one of the most important life characteristics of humans. Tumor cells proliferate more rapidly and increase in number geometrically compared to normal cells. It also destroys the life cycle and division cycle of normal cells and acquires the ability to prolong proliferation and growth, which is one of the difficulties in treating cancer [25, 26].

Methylation-driven Drug Gene Association Network

We used the DD Score method to re-target methylated human FLNs to obtain a methylation-driven drug gene association network. Literature mining for breast cancer resetting drug targets and resetting drugs in the network showed that a large number of studies have reported that breast cancer reset drug targets are closely related to cancer, and may serve as novel drugs for breast cancer treatment. For example, López-Ozuna *et al.*[27] found that the mRNA and protein expression of the prolactin receptor (PRLR) is downregulated in triple-negative breast cancer (TNBC). At the same time, PRLR gene expression has been used to subclassify TNBC patients into a new subgroup (TNBC-PRLR) characterized by epithelial luminal differentiation. The restoration/activation of the PRL pathway in TNBC cells representative of mesenchymal or TNBC-PRLR subgroups led to induction of the epithelial phenotype and suppression of tumorigenesis. These results offer potential new modalities for TNBC stratification and development of personalized therapy based on PRL pathway activation. Kelly [28] found that PRLR is expressed in breast cancer and promotes cancer. Urban studies [29] have found that the mRNA and protein expression of glycoprotein M6B (GPM6B) is associated with various types of gynecologic malignancies. N-myc downstream regulatory gene 2 (NDRG2), a well-known tumor suppressor, was identified as a potential target gene for miR-454, and miR-454 directly targets the 3'-untranslated region of NDRG2. Gao [30] found that miR-454 is a key regulator of tumorigenesis, and can act as an oncogene or a tumor suppressor. At the same time, they found that miR-454 overexpression significantly inhibited NDRG2 expression, and small interfering RNA (siRNA)-mediated NDRG2 significantly reversed the anti-tumor effect of miR-454 inhibition on cancer cell proliferation and invasion. Kim [31] found that inhibition of NF- κ B signaling by NDRG2 gene expression downregulated cyclooxygenase-2 expression, thereby inhibiting the migration and invasion of breast cancer cells (Fig. 4A).

Copy Number-driven Drug Gene Association Network

We used the DD Score method to re-target the copy number human FLNs to obtain a copy number-driven drug gene association network. For example, Yen [32] found that expression of acyl-CoA synthetase long chain family member 1 (ACSL1) in the estrogen receptor (ER)-negative group was higher than that in the ER-positive group, and high ACSL5 expression was associated with good prognosis in patients with both ER-positive and ER-negative breast cancer. These results suggest that ACSL1 expression is regulated by ER signaling pathways, and may be a potential novel biomarker for predicting the prognosis of breast cancer patients. Wang [33] found that ACSL1 mRNA expression levels were significantly upregulated in clinical breast cancer tissues, and oncoprotein Hepatitis B virus X-interacting protein upregulated ACSL1 in breast cancer by activating transcription factor Sp1. The Aikins study [34] found that transgelin is an actin-binding protein associated with multiple stages of cancer development, such as proliferation, migration, and invasion. The authors studied the role of transgelin in vasculogenic mimicry (VM) and evaluated the effects of VM on endothelial cell- and angiogenesis related-gene expression in MDA-MB-231 breast cancer cells. The results suggested that downregulation of transgelin (TAGLN) inhibits interleukin-8, thereby inhibiting VM in breast cancer cells, and through this mechanism, TAGLN and VM together inhibit malignant tumors. Many studies have shown that the high expression of Kinesin Family Member 26B (KIF26B) is directly related to a poor prognosis in breast cancer. Teng's study [35] found a significant increase in KIF26B levels in breast cancer cells and patient samples, and KIF26B levels correlated with tumor size, TNM grade, and degree of differentiation in breast cancer patients. Overexpression of KIF26B promotes breast cancer cell proliferation and migration by activating FGF2/ERK signaling, whereas silencing KIF26B has the opposite effect. Experimental studies have shown that KIF26B promotes the development and progression of breast cancer and may become a potential therapeutic target for the treatment of breast cancer. Non-receptor tyrosine kinases, such as protein tyrosine kinase 6 (PTK6), are highly expressed in a variety of tumor types, including prostate, ovarian, and breast cancers, and modulate oncogenic phenotypes such as proliferation, migration, and survival. PTK6 inhibition also overcomes targeted therapy resistance of human epidermal growth factor 2-positive breast cancer. Although PTK6 is highly expressed in ER + Luminal breast cancers, the role of PTK6 in this subtype has not been elucidated. The Park study [36] elucidated the critical functions of PTK6 in ER + Luminal breast cancers and support PTK6 as an attractive therapeutic target for ER + breast cancer (Fig. 4B).

Mutation-driven Drug Gene Association Network

We used the DD Score method to re-target the mutation human FLNs to obtain a mutation-driven drug gene association network. For example, Litviakov's experimental study [37] indicated that the absence of ABCC5 predicted a good response to neoadjuvant chemotherapy in breast cancer. Li's experimental study [38] found that high Minichromosome Maintenance 10 Replication Initiation Factor (MCM10) immune expression was significantly associated with the presence of advanced primary tumors, lymph node status, and vascular invasion in cancer. Low expression of MCM10 gene significantly inhibited cell

proliferation in J82 and TCCSUP human bladder carcinoma cells. The experimental results indicated that MCM10 overexpression was associated with poor clinical pathological features and independent negative prognosis, which have potential therapeutic diagnostic value in cancer. Fu [39] found that solute carrier family 22 member 3 (SLC22A3) is a metastasis suppressor in esophageal squamous cell carcinoma (ESCC), which is significantly downregulated in non-tumor esophageal tissues of patients with familial ESCC and is significantly associated with lymph node metastasis. A-to-I RNA editing of the SLC22A3 gene results in reduced expression in non-tumor esophageal tissues of familial ESCC. Their study showed that A-to-I RNA editing of SLC22A3 contributed to the early development and progression of familial esophageal cancer in high-risk individuals (Fig. 4C).

Multiomics Joint-driven Drug Gene Association Network

We used the DD Score method to retarget the multi-omics human functional linkages network to obtain a multi-omics joint-driven drug gene association network. For example, placental-specific protein 1 (Plac1) is a cancer antigen that plays a key role in promoting the initiation and progression of cancer. Feng [40] study found that Plac1 expression was positively correlated with clinical stage, lymph node metastasis, hormone receptor status, and overall patient survival. Overexpression of Plac1 promoted invasion and metastasis of breast cancer cells *in vitro* and *in vivo*. The experimental results indicate that the functional interaction between Plac1 and furin enhances the invasion and metastasis of breast cancer, and the furin/NICD/PTEN axis may be an important therapeutic target for breast cancer treatment. Yang [41] study found that Plac1 is a cancer-associated protein that is upregulated in a variety of malignant tissues including prostate cancer, gastric adenocarcinoma, colorectal cancer, epithelial ovarian cancer, and breast cancer. It is significantly upregulated in cancer tissues, and its expression level is associated with advanced pathological stage and shorter progression-free survival of cancer patients. Furthermore, downregulation of Plac1 expression by siRNA inhibited cell proliferation, induced apoptosis, and disrupted the invasive ability of cancer cells. These results suggest that upregulated Plac1 can be used as a negative prognostic biomarker in cancer and regulate cell proliferation and invasion. Eriksson [42] found that breast size-related gene (COMTD1) mutations affect breast cancer risk. FAM107A (TU3A), located at 3p21.1, was identified as a candidate tumor suppressor gene in renal cell carcinoma. Awakura [43] found that TU3A promoter hypermethylation was observed in several types of cancer cell lines and primary cancers of the bladder and testis. These results demonstrated for the first time the epigenetic inactivation of TU3A in human cancers and have important implications for studying the role of TU3A methylation in cancer development. Also known as Gravin and SSeCKS, AKAP12 is a novel potent scaffolding protein for many key signaling factors, such as protein kinase C (PKC), PKA, cyclin and F-actin. AKAP12 expression is inhibited in a variety of malignancies including breast cancer, prostate cancer, gastric cancer, and colon cancer. Soh [44] found that AKAP12 is a potential breast cancer metastasis suppressor. Zhang [45] found that the non-methylation of AOX1 is a risk factor for breast cancer.

Prognostic Utility Of Reset Drug Targets In Breast Cancer

The DD Score method was used to reset drug targets for breast cancer, and the survival and survminer R packages were used to perform survival analysis. The targets significantly classified high-risk and low-risk groups of breast cancer patients with strong prognostic efficacy. The methylation-driven gene GPM6B has a P value of 0.00141, a copy number driven gene ACSL1, a P value of 0.0582, a mutation-driven gene MCM10, a P value of 0.0303, and a multiomics joint-driven drug gene COMTD1 with a P value of 0.0071. A significantly high expression of GPM6B, a methylation-driven breast cancer replacement drug target, was significantly associated with a better prognosis. Urban studies have shown that increased mRNA and protein expression of Gpm6B is associated with various types of gynecologic malignancies[29]. At the same time, its targeted drug methotrexate is an anti-tumor drug with immunosuppressive properties, which can be used alone or in combination with other anticancer agents for the treatment of breast cancer. Significant low expression of ACSL1 is markedly associated with a better prognosis in copy-driven breast cancer replacement drug targets. Yen found that ACSL1 expression is regulated by the ER signaling pathway and is a potential novel biomarker for predicting the prognosis of breast cancer patients[32]. At the same time, its targeted drug mitomycin is an anti-tumor antibiotic produced by *Streptomyces caespitosus* for the treatment of malignant tumors of the lips, mouth, pharynx, digestive organs, peritoneum, female breasts and bladder. Significantly low expression of MCM10, a mutation-driven breast cancer replacement drug target, is significantly associated with better prognosis. Li found that high MCM10 immune expression is significantly associated with the presence of advanced primary tumors, lymph node status, and vascular invasion in cancer[38]. Low gene expression of MCM10 significantly inhibited cell proliferation in J82 and TCCSUP cells. Its targeted drugs including Mitomycin, Methotrexate and Estradiol. Estradiol is a naturally occurring hormone that can be used to treat breast cancer (for relief only). Significantly low expression of the breast cancer replacement drug target COMTD1, a multiomics combination-driven drug gene network, was significantly associated with better prognosis. The Eriksson N study found that breast size-related gene (COMTD1) mutations affect breast cancer risk[42]. Its targeted drug, cisplatin, is a platinum-based chemotherapy drug used to treat various types of cancer, including sarcoma, small cell lung cancer, ovarian cancer, lymphoma, and germ cell tumors.

Discussion

We used breast cancer as a disease model to systematically propose a new strategy based on reconstruction of human FLN drug repositioning. Through pathway enrichment analysis, it was found that DNA methylation, copy number and mutation disease risk genes were significantly enriched in "PI3K-Akt signaling pathway", "pathways in cancer", "CAMs," and "miRNAs in cancer". Through functional enrichment analysis, it was found that DNA methylation, copy number, and mutation disease risk genes were significantly enriched in the "angiogenesis, negative regulation of apoptotic process", "positive regulation of cell proliferation/negative regulation of cell proliferation," and "angiogenesis." For methylation, copy number, mutation, and multiomics to drive networks, literature mining showed that our

breast cancer reset drug targets and replacement drugs are strongly associated with breast cancer. TNBC is currently the most malignant breast cancer subtype with no effective targeted therapy. Mifepristone (MIF) is a progestogen and glucocorticoid antagonist. By releasing endogenous prostaglandins from the endometrium or decidua, their inhibition of progesterone induces bleeding during the luteal phase and early pregnancy. As a glucocorticoid receptor antagonist, the drug is used as a drug for abortion and also for the treatment of hypercortisolism in patients with non-pituitary Cushing's syndrome. It has anti-tumor activity in a variety of hormone-dependent cancers, including luminal breast cancer. Liux perimentally found that MIF inhibited tumor growth of TNBC cell lines and patient-derived xenografts in NOD-SCID mice[46]. Furthermore, MIF reduces TNBC cancer stem cell (CSC) population by down-regulating KLF5 expression, a stem cell transcription factor overexpressing in basal TNBC and promoting cell proliferation. MIF inhibits the expression of KLF5 by inducing expression of miR-153. Consistently, miR-153 reduced CSC and miR-153 inhibitors to rescue MIF-induced downregulation of KLF5 protein levels and CSC ratios. Experimental results indicate that MIF inhibits basal TNBC via the miR-153/KLF5 axis, and MIF can be used to treat TNBC. Multiple drug resistant (MDR) malignancy remains a predictable and often terminal event in cancer therapy, and affects individuals with many cancer types, regardless of the stage at which they were originally diagnosed or the interval from last treatment. Protein biomarkers of MDR are not globally used for clinical decision-making, but they include the overexpression of drug-efflux pumps (ABC transporter family), such as MDR-1 and BCRP, as well as HIF1 α , a stress responsive transcription factor found elevated within many MDR tumors. Metformin is considered an anti hyperglycemic drug because it lowers blood glucose concentrations in type II diabetes without causing hypoglycemia. Metformin is commonly described as an insulin sensitizer leading to a decrease in insulin resistance and a clinically significant reduction of plasma fasting insulin levels. Another well-known benefit of this drug is modest weight loss. Metformin is the drug of choice for obese type II diabetes patients. Davies G found that metformin, at low physiological concentrations, reduces the expression of multiple classic protein markers of MDR *in vitro* and in preliminary *in vivo* models. The experiment demonstrates the effectiveness of metformin in resensitizing MDR breast cancer cell lines to their original treatment, and it provides evidence that metformin may function through a mechanism involving post-translational histone modifications via an indirect histone deacetylase inhibitor (HDACi) activity. The results demonstrate that metformin can prevent MDR development and resensitize MDR cells to chemotherapy *in vitro*, and metformin may be a potential therapeutic drug for the treatment of MDR breast cancer. This provides important medical relevance towards metformin's potential clinical use against MDR cancers[47]. Urokinase-type plasminogen activator (uPA) has been validated as a predictive or prognostic biomarker protein, and mesupron is considered the first-in-class anticancer agent to inhibit uPA activity in human breast cancer. Auranofin is a gold salt that is capable of eliciting pharmacologic actions that suppress inflammation and stimulate cell-mediated immunity. Auranofin appears to induce heme oxygenase 1 mRNA. Heme oxygenase 1 is an inducible heme-degrading enzyme with anti-inflammatory properties. It is therefore listed by the World Health Organization as a member of the anti-rheumatic agent category. Lee found that synergy between mesupron and auranofin can be used to induce apoptosis in MCF-7 human breast cancer cells[48]. The results demonstrated that mesupron and auranofin significantly lead to inhibition of cancer cell proliferation; cell cycle arrest at the G1/S phase of

the cell cycle; and apoptosis as indicated by caspase 3 activation, poly(ADP-ribose) polymerase cleavage, and annexin V staining. This combined treatment decreased the levels of mitochondrial anti-apoptotic factors, such as BCL-2, BCL-xL, and MCL-1 and caused nuclear translocation of apoptosis-inducing factor. Mitochondrial membrane potential was found to be strongly disrupted in combination-treated cells. In addition, combination treatment significantly enhanced the overproduction of reactive oxygen species, which was rescued by N-acetylcysteine treatment. The combination treatment suppressed phosphorylation of Akt, thus contributing to apoptosis. Experimental studies have shown that the use of mesupron in combination with auranofin may be important for achieving high anticancer synergy, which may be a potential combination therapy for the treatment of breast cancer. Through survival analysis, we found that targets can significantly classify high-risk and low-risk groups of breast cancer patients with strong prognostic efficacy. For example, the methylation gene GPM6B (P = 0.00141), the copy number gene ACSL1 (P = 0.0582), the mutant gene MCM10 (P = 0.0303), and the multi-omics data gene COMTD1 (P = 0.0071). The results show that our method can well identify breast cancer strong related genes and potential breast cancer treatment drugs, provide important medical relevance for clinical research of breast cancer drugs, and provide more effective drug discovery for drug relocation research channel.

Conclusions

In this study, we used breast cancer as a disease model to systematically propose a new strategy for drug repositioning and to identify drug-associated biomarkers based on reconstruction of the human functional linkage network (FLN). The goal of the study was to find disease risk genes that have strong functional relationships with breast cancer drug target genes, in order to reconstruct human FLNs for drug relocation. In particular, we searched for drugs that strongly perturb genes with the following properties: strong functional connections; abnormal changes in the molecule during the disease stage (e.g., mutation, DNA copy number variation, DNA methylation); expression levels that are strongly interfered with by the disease; and significant enrichment in biological pathways affected by disease. Our system interpreted functional relationships among human genes and established a measure of the relationship between drugs and diseases to identify drug-based prognostic biomarkers.

Abbreviations

lncRNA: long non-coding RNA, BRCA: breast cancer, PCG: protein coding gene, FDA: Food and Drug Administration, DEGs: differentially expressed genes, FLN: human functional linkage network,

Declarations

Ethics approval and consent to participate

Patient data we used were acquired by publicly available datasets that were collected with patients' informed consent.

Consent for publication

All authors agreed on the manuscript.

Availability of supporting data

Multi-omics data of breast cancer and survival data of patients are available in TCGA Data Portal. FDA-approved drugs and their target genes were obtained from Drug Bank and human FLN was obtained from the visANT database

Competing Interests

No potential conflicts of interest were disclosed.

Funding

This work was financially supported by National Natural Science Foundation of China (31801115,61803129), the China Postdoctoral Science Foundation (2018M631943, 2018M641860), the China Postdoctoral Science Special Foundation (2019T120280), the Hei Long Jiang Postdoctoral Special Foundation (LBH-TZ1018), the Hei Long Jiang Postdoctoral Foundation (LBH-Z17110, LBH-Z17218), and Heilongjiang Provincial Health and Family Planning Commission of Science Foundation (2018476,2018477). Heilongjiang Provincial planning office key subjects (GBB1318066).

Authors' contributions

HYZ, LY and YYP implemented the algorithm, analyzed the data and wrote the manuscript. LW, SWN and JL conceived and designed the study. SHL, CYZ, ZJL, JJL, HTX, TYL and WDH collected the data and helped to write the paper. All authors read and approved the final manuscript.

Acknowledgements

The TCGA datasets that support the findings of this study.

Authors' information

1 College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China. 2 Department of Ultrasonic medicine, The 1st Affiliated Hospital of Heilongjiang University of Chinese Medicine, Harbin 150040, China.

References

1. Chong CR, Sullivan DJ Jr. New uses for old drugs. *Nature*. 2007;448:645–6.
2. DiMasi JA, Hansen RW, Grabowski HG. The price of innovation: new estimates of drug development costs. *J Health Econ*. 2003;22:151–85.

3. Renaud RC, Xuereb H. Erectile-dysfunction therapies. *Nat Rev Drug Discov.* 2002;1:663–4.
4. Lin TS, Prusoff WH. Synthesis and biological activity of several amino analogues of thymidine. *J Med Chem.* 1978;21:109–12.
5. Shaughnessy AF. Old drugs, new tricks. *BMJ.* 2011;342:d741.
6. Khan SA, Virtanen S, Kallioniemi OP, Wennerberg K, Poso A, Kaski S. Identification of structural features in chemicals associated with cancer drug response: a systematic data-driven analysis. *Bioinformatics.* 2014;30:i497–504.
7. Li P, Huang C, Fu Y, Wang J, Wu Z, Ru J, Zheng C, Guo Z, Chen X, Zhou W, et al. Large-scale exploration and analysis of drug combinations. *Bioinformatics.* 2015;31:2007–16.
8. Keiser MJ, Setola V, Irwin JJ, Laggner C, Abbas AI, Hufeisen SJ, Jensen NH, Kuijter MB, Matos RC, Tran TB, et al. Predicting new molecular targets for known drugs. *Nature.* 2009;462:175–81.
9. Campillos M, Kuhn M, Gavin AC, Jensen LJ, Bork P. Drug target identification using side-effect similarity. *Science.* 2008;321:263–6.
10. Iorio F, Bosotti R, Scacheri E, Belcastro V, Mithbaekar P, Ferriero R, Murino L, Tagliaferri R, Brunetti-Pierri N, Isacchi A, di Bernardo D. Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc Natl Acad Sci U S A.* 2010;107:14621–6.
11. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res.* 2005;33:D514–7.
12. Rogers FB. Medical subject headings. *Bull Med Libr Assoc.* 1963;51:114–6.
13. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science.* 2006;313:1929–35.
14. Linghu B, Snitkin ES, Hu Z, Xia Y, Delisi C. Genome-wide prioritization of disease genes and identification of disease-disease associations from an integrated human functional linkage network. *Genome Biol.* 2009;10:R91.
15. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhi R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* 2011;12:R41.
16. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, Kong L, Gao G, Li CY, Wei L. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* 2011;39:W316–22.
17. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4:44–57.
18. Chen HR, Sherr DH, Hu Z, DeLisi C. A network based approach to drug repositioning identifies plausible candidates for breast cancer and prostate cancer. *BMC Med Genomics.* 2016;9:51.
19. Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. *Cell.* 2010;141:1117–34.

20. Osaki M, Oshimura M, Ito H. PI3K-Akt pathway: its functions and alterations in human cancer. *Apoptosis*. 2004;9:667–76.
21. Hers I, Vincent EE, Tavare JM. Akt signalling in health and disease. *Cell Signal*. 2011;23:1515–27.
22. Dejana E. Endothelial cell-cell junctions: happy together. *Nat Rev Mol Cell Biol*. 2004;5:261–70.
23. Krauss RS, Cole F, Gaio U, Takaesu G, Zhang W, Kang JS. Close encounters: regulation of vertebrate skeletal myogenesis by cell-cell contact. *J Cell Sci*. 2005;118:2355–62.
24. de Mendez MT, Bosch AL. Abnormal immunoexpression of cell adhesion molecules (CAMs) in cervical cancer. *Int J Surg Pathol*. 2011;19:733–42.
25. Lyu J, Yang EJ, Shim JS. **Cholesterol Trafficking: An Emerging Therapeutic Target for Angiogenesis and Cancer**. *Cells* 2019, 8.
26. Gu S, Xue J, Xi Y, Tang R, Jin W, Chen JJ, Zhang X, Shao ZM, Wu J. Evaluating the effect of Avastin on breast cancer angiogenesis using synchrotron radiation. *Quant Imaging Med Surg*. 2019;9:418–26.
27. Lopez-Ozuna VM, Hachim IY, Hachim MY, Lebrun JJ, Ali S. Prolactin Pro-Differentiation Pathway in Triple Negative Breast Cancer: Impact on Prognosis and Potential Therapy. *Sci Rep*. 2016;6:30934.
28. Kelly MP, Hickey C, Makonnen S, Coetzee S, Jalal S, Wang Y, Delfino F, Shan J, Potocky TB, Chatterjee I, et al. Preclinical Activity of the Novel Anti-Prolactin Receptor (PRLR) Antibody-Drug Conjugate REGN2878-DM1 in PRLR-Positive Breast Cancers. *Mol Cancer Ther*. 2017;16:1299–311.
29. Urban P, Rabajdova MB, Varga J, Kruzliak P, Fialkovicova V, Durovcova E, Marekova M. Vascular marker expression during the development of various types of gynaecological malignancy. *Tumour Biol*. 2014;35:11229–35.
30. Fu Q, Gao Y, Yang F, Mao T, Sun Z, Wang H, Song B, Li X. Suppression of microRNA-454 impedes the proliferation and invasion of prostate cancer cells by promoting N-myc downstream-regulated gene 2 and inhibiting WNT/beta-catenin signaling. *Biomed Pharmacother*. 2018;97:120–7.
31. Kim MJ, Kim HS, Lee SH, Yang Y, Lee MS, Lim JS. NDRG2 controls COX-2/PGE(2)-mediated breast cancer cell migration and invasion. *Mol Cells*. 2014;37:759–65.
32. Yen MC, Kan JY, Hsieh CJ, Kuo PL, Hou MF, Hsu YL. Association of long-chain acyl-coenzyme A synthetase 5 expression in human breast cancer by estrogen receptor status and its clinical significance. *Oncol Rep*. 2017;37:3253–60.
33. Wang Y, Cai X, Zhang S, Cui M, Liu F, Sun B, Zhang W, Zhang X, Ye L. HBXIP up-regulates ACSL1 through activating transcriptional factor Sp1 in breast cancer. *Biochem Biophys Res Commun*. 2017;484:565–71.
34. Aikins AR, Kim M, Raymundo B, Kim CW. Downregulation of transgelin blocks interleukin-8 utilization and suppresses vasculogenic mimicry in breast cancer cells. *Exp Biol Med (Maywood)*. 2017;242:573–83.
35. Teng Y, Guo B, Mu X, Liu S. KIF26B promotes cell proliferation and migration through the FGF2/ERK signaling pathway in breast cancer. *Biomed Pharmacother*. 2018;108:766–73.

36. Ito K, Park SH, Katsyv I, Zhang W, De Angelis C, Schiff R, Irie HY. PTK6 regulates growth and survival of endocrine therapy-resistant ER + breast cancer cells. *NPJ Breast Cancer*. 2017;3:45.
37. Litviakov NV, Cherdyntseva NV, Tsyganov MM, Slonimskaya EM, Ibragimova MK, Kazantseva PV, Kzhyshkowska J, Choinzonov EL. Deletions of multidrug resistance gene loci in breast cancer leads to the down-regulation of its expression and predict tumor response to neoadjuvant chemotherapy. *Oncotarget*. 2016;7:7829–41.
38. Li WM, Huang CN, Ke HL, Li CC, Wei YC, Yeh HC, Chang LL, Huang CH, Liang PI, Yeh BW, et al. MCM10 overexpression implicates adverse prognosis in urothelial carcinoma. *Oncotarget*. 2016;7:77777–92.
39. Fu L, Qin YR, Ming XY, Zuo XB, Diao YW, Zhang LY, Ai J, Liu BL, Huang TX, Cao TT, et al. RNA editing of SLC22A3 drives early tumor invasion and metastasis in familial esophageal cancer. *Proc Natl Acad Sci U S A*. 2017;114:E4631–40.
40. Li Y, Chu J, Li J, Feng W, Yang F, Wang Y, Zhang Y, Sun C, Yang M, Vasilatos SN, et al. Cancer/testis antigen-Plac1 promotes invasion and metastasis of breast cancer through Furin/NICD/PTEN signaling pathway. *Mol Oncol*. 2018;12:1233–48.
41. Yang L, Zha TQ, He X, Chen L, Zhu Q, Wu WB, Nie FQ, Wang Q, Zang CS, Zhang ML, et al. Placenta-specific protein 1 promotes cell proliferation and invasion in non-small cell lung cancer. *Oncol Rep*. 2018;39:53–60.
42. Eriksson N, Benton GM, Do CB, Kiefer AK, Mountain JL, Hinds DA, Francke U, Tung JY. Genetic variants associated with breast size also influence breast cancer risk. *BMC Med Genet*. 2012;13:53.
43. Awakura Y, Nakamura E, Ito N, Kamoto T, Ogawa O. Methylation-associated silencing of TU3A in human cancers. *Int J Oncol*. 2008;33:893–9.
44. Soh RYZ, Lim JP, Samy RP, Chua PJ, Bay BH. A-kinase anchor protein 12 (AKAP12) inhibits cell migration in breast cancer. *Exp Mol Pathol*. 2018;105:364–70.
45. Zhang H, Liu YP, Ge AQ, Wang X, Sun HR, Bi HR, Pang D, Zhao YS. [Association between AOX1, IRF4 methylation in peripheral blood leukocyte DNA and the risks of breast cancer: a case-control study]. *Zhonghua Liu Xing Bing Xue Za Zhi*. 2018;39:1265–9.
46. Liu R, Shi P, Nie Z, Liang H, Zhou Z, Chen W, Chen H, Dong C, Yang R, Liu S, Chen C. Mifepristone Suppresses Basal Triple-Negative Breast Cancer Stem Cells by Down-regulating KLF5 Expression. *Theranostics*. 2016;6:533–44.
47. Davies G, Lobanova L, Dawicki W, Groot G, Gordon JR, Bowen M, Harkness T, Arnason T. Metformin inhibits the development, and promotes the resensitization, of treatment-resistant breast cancer. *PLoS One*. 2017;12:e0187191.
48. Lee JE, Kwon YJ, Baek HS, Ye DJ, Cho E, Choi HK, Oh KS, Chun YJ. Synergistic induction of apoptosis by combination treatment with mesupron and auranofin in human breast cancer cells. *Arch Pharm Res*. 2017;40:746–59.

Figures

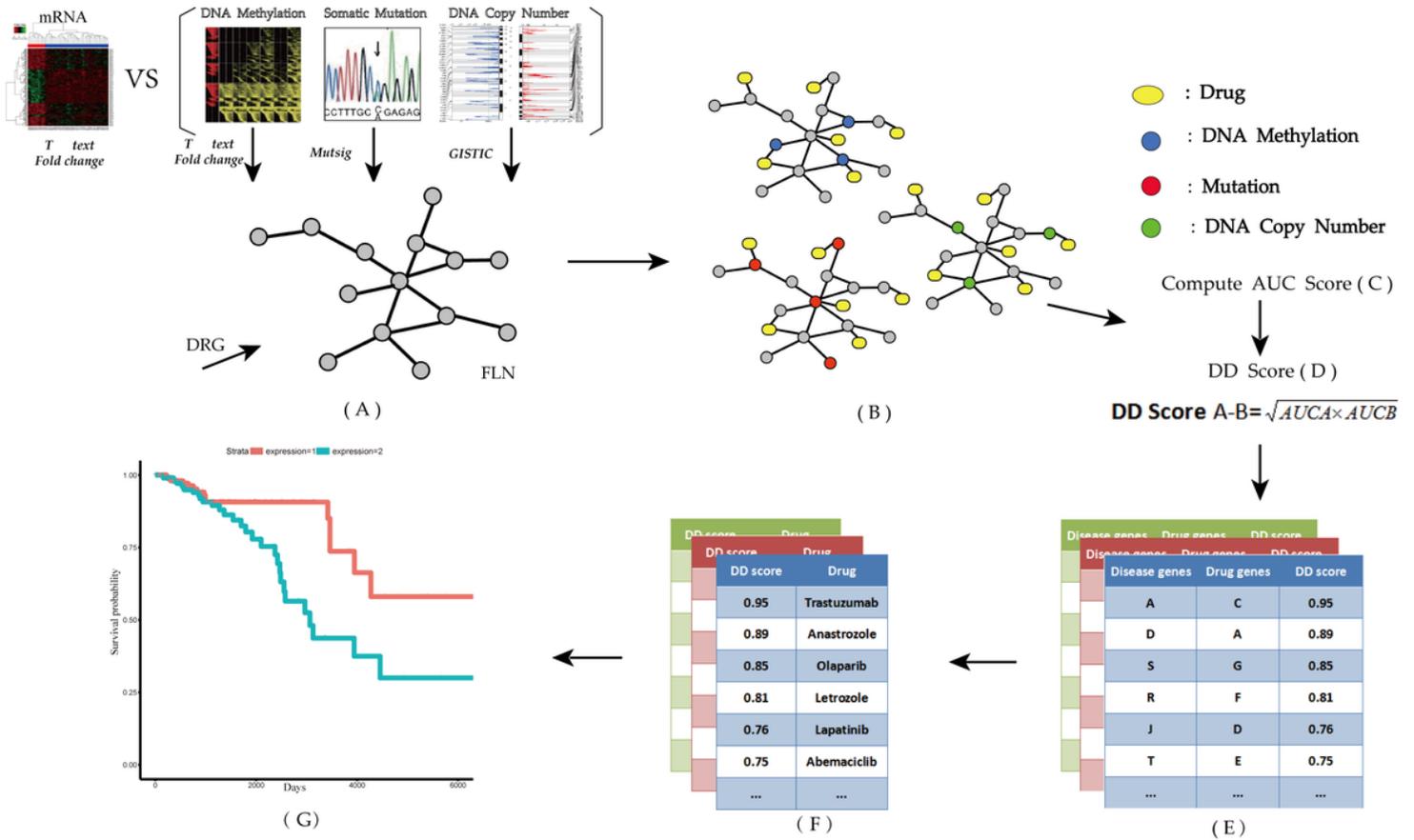


Figure 1

Flowchart of the strategy to identify drug-based prognostic biomarkers. (A) Mapping the methylation, copy number, and mutation risk disease genes to the FLN respectively, and mapping DRG to the FLN. (B) Reconstructing the human functional connection networks linkage network. (C) Computing the AUC value. (D). Computing the DD Scores. (E). Sorting the gene pairs according to the DD score, and literature mining of top 100 gene pairs. (F). According to the literature mining and DD scores to get repositioning drugs. (G). Repositioning drug target survival analysis processing.

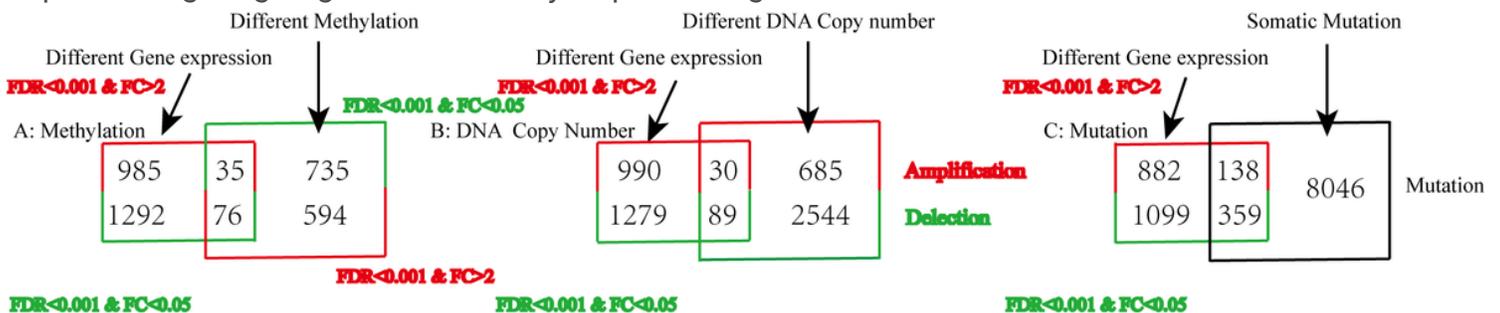


Figure 2

Identification of breast cancer risk genes based on multiomics (A) Among the methylated DEGs, red is hypermethylated and has low expression, and green is hypomethylated and is highly expressed. (B) Among the copy number difference genes, red is copy-amplified and is highly expressed, and green is

copy-missed and has low expression. (C) Among the mutation differential gene, black is mutation gene. DNA methylation, DNA copy number, and somatic mutations are consistent with mRNA expression.

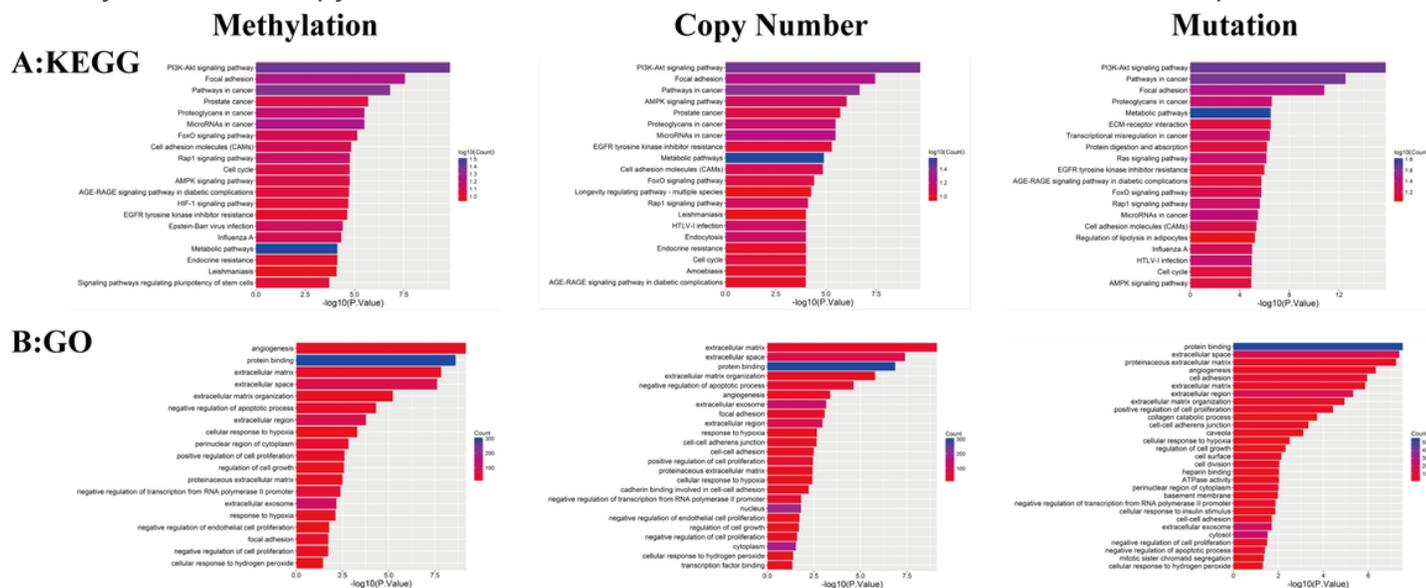


Figure 3

Identification of biological pathways and functions in breast cancer. (A) Methylation, copy number, and mutation with significantly enriched pathways. (B)Methylation, copy number, and mutation with significantly enrichment functions.

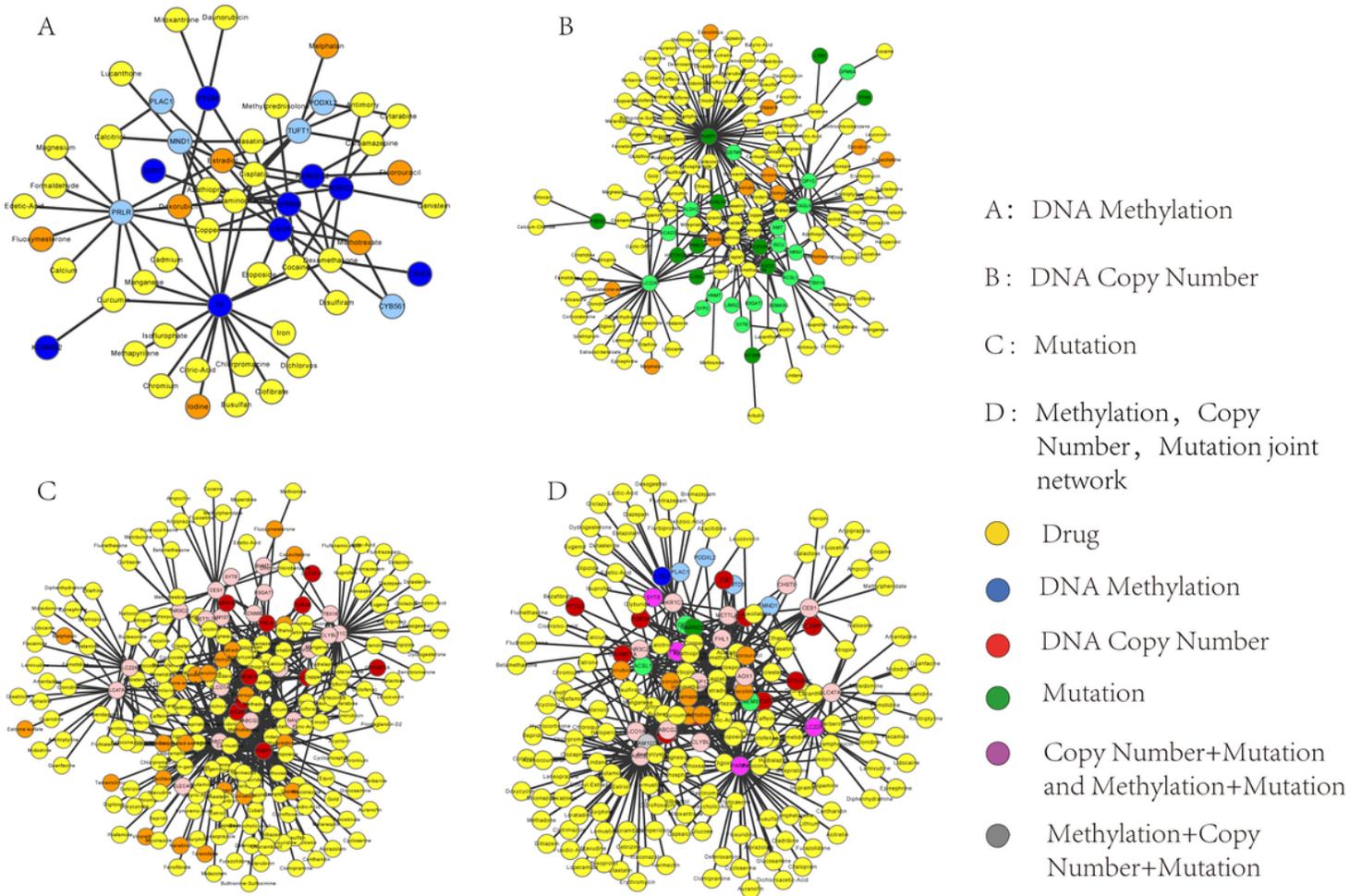


Figure 4

Reconstruction of human function linkages network. (A) Methylation-driven drugs-genes association network. (B) Copy number-driven drug-genes association network. (C) Mutation-driven drug-genes association network. (D) Multiomics joint-driven drugs-genes association network. In the whole picture, dark yellow represents breast cancer drugs, and light yellow represents non-breast cancer drugs. Dark blue represents gene that is low expression by hypermethylation, and light blue represents gene that is high expression by hypomethylation. Dark green represents gene that is high expressed by copy number amplification, and light green represents gene with low expression of copy number missing. Dark red represents gene with high expression of mutation, and light red represents a gene with low expression of mutation. Purple represents copy number and mutation or methylation and mutation intersection gene. Gray represents the intersection gene of methylation, copy number and mutation.

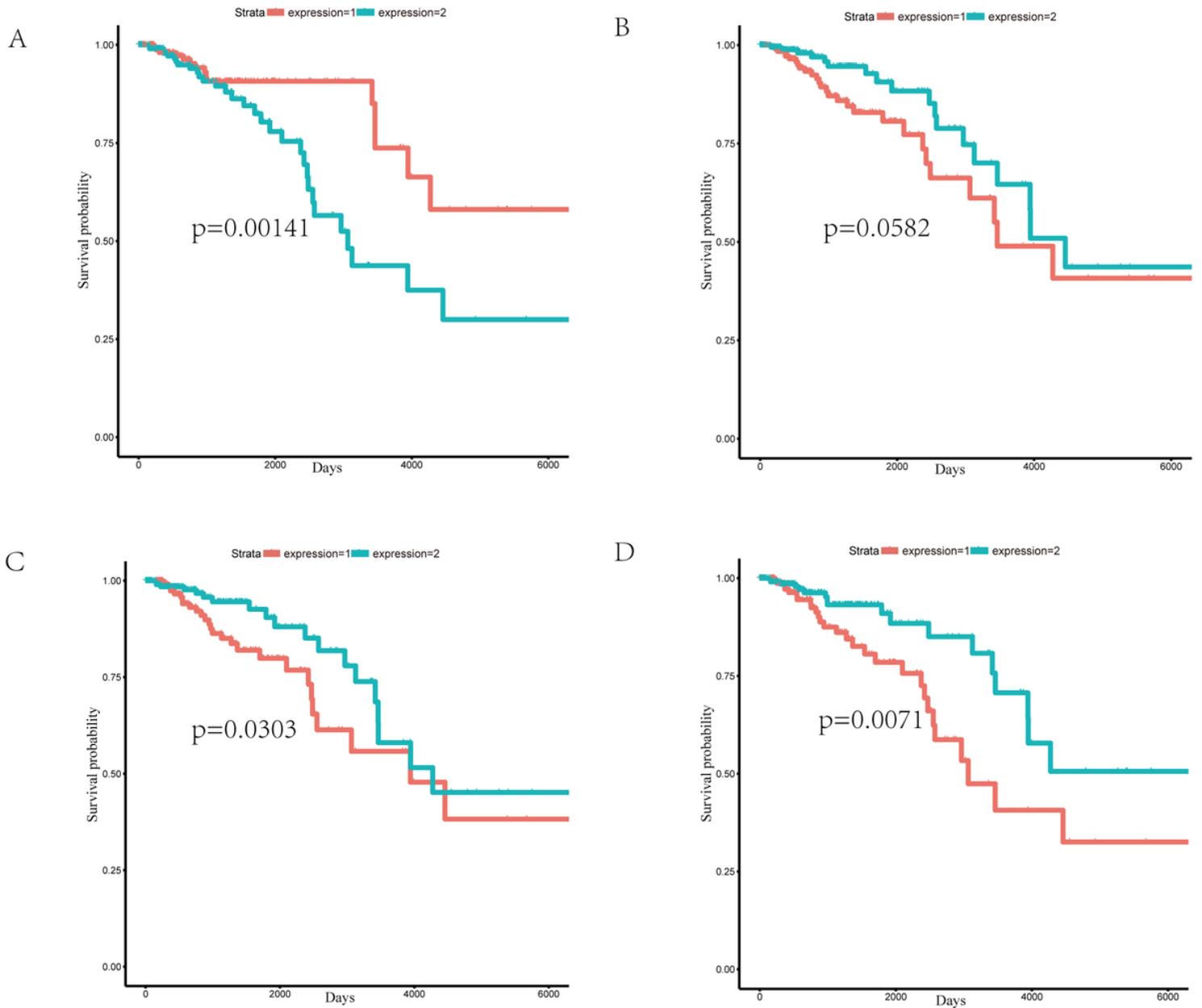


Figure 5

Correlation of drug-related biomarkers with overall survival (OS) of Breast cancer. (A) Methylation-driven gene (GPM6B). (B) Copy number-driven gene (ACSL1). (C). Mutation-driven gene (MCM10). (D) Methylation, copy number, mutation combined-driven gene (COMTD1).