Article

J. Biol. Regul. Homeost. Agents. 2023; 37(4): 2267–2276
https://doi.org/10.23812/j.biol.regul.homeost.agents.20233704.223

Identification of Key Genes and Pathways Associated with Adriamycin Resistance in Breast Cancer

Ting Liu[1], Yan Xue[1], Ping-Ping Guo[1], Qi Wang[1], Shu Lin[1,*]

[1]School of Pharmacy, Shenyang Medical College, 110034 Shenyang, Liaoning, China
*Correspondence: linshulsnm0@126.com (Shu Lin)
Published: 20 April 2023

Objective: Drug resistance to the chemotherapeutic drug Adriamycin (ADR) is a key clinical impediment to successful breast cancer treatment (BrCa). However, the molecular mechanism and targets that mediate ADR resistance remain unclear. Therefore, the identification of ADR response biomarkers to improve the treatment of patients with BrCa is an urgent issue.
Methods: The GSE24460 dataset on Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) was employed in this investigation, which covers gene expression profiles of parental and ADR-resistant cell lines of MCF-7 (cell lines in the human breast cancer). Differentially expressed genes (DEGs) in parental and ADR-resistant cells, Gene Ontology (GO) enrichment analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway studies, and the interactions of protein-protein were used to identify the overall biological alterations (PPIs). The most prevalent DEGs in the PPI, GO, and KEGG pathways were discovered and using GSE34138 dataset from the GEO database were validated, and their ability to forecast overall survival (OS) and relapse-free survival (RFS) was clinically validated. The link between the DEGs and OS/RFS was studied after further verification of the most common DEGs in those pathways. The most significant crucial gene, collagen type-IV alpha 1 (*COL4A1*), was discovered in the MCF-7 and MCF-7/ADR cells using the quantitative real-time reverse transcription PCR, Western blot, and Cell Counting Kit-8 assays to further highlight the distinctions.
Results: 207 DEGs in total were discovered, with 111 upregulated and 96 downregulated. After mapping the GOs, pathways, and PPI networks, six genes—cyclin-dependent kinase inhibitor 2A (*CDKN2A*), retinoblastoma 1 (*RB1*), C-X-C motif ligand 12 (*CXCL12*), *COL4A1*, intercellular cell molecule-1 of adhesion (*ICAM1*), and cadherin 1 (*CDH1*)—were found to be frequently enriched. *COL4A1* was found to be favorably associated with ADR resistance and poor OS, as well as strongly associated with RFS. The correlation between *CDH1* and ADR resistance was negative, which was substantial in terms of poor RFS. In MCF-7/ADR cells, *COL4A1* also expression was noted to have increased, and *COL4A1* knockdown significantly reduced the inhibitory concentration 50% value of ADR in those cells.
Conclusions: Our data identified the significant pathways and genes for predicting the emergence of ADR resistance and revealed that *COL4A1* regulated the proliferation of MCF-7/ADR and played a crucial role in ADR resistance. Therefore, *COL4A1* shows that it can be novel target for improving patients' prognosis with ADR-resistant BrCa.

Keywords: KEGG; PPI; Adriamycin resistance; breast cancer; key genes; pathways

Introduction

In women worldwide, the most frequent malignant tumor is Breast cancer (BrCa), it accounts for 31% (287,850) of all newly diagnosed malignancies in the United States by 2021 [1]. Breast cancer has the highest general malignancy of all cancers, with serious impacts on patients' survival and quality of life [1]. For those with advanced BrCa, the most common treatment strategy is surgical resection combined with adjuvant chemotherapy. However, the curative effect of chemotherapy is often weakened by the occurrence of drug resistance [2]. The widely used agent is Adriamycin (ADR) in the treatment chemotherapy of BrCa and is significantly proven to enhance the prognosis of patients with BrCa [3]. However, the key barrier is ADR to the clinical treatment of BrCa is successful manner, resulting in subsequent tumor development and recurrence, with eventual therapeutic failure [4]. The mechanisms of ADR resistance are quite complex and remain far from fully understood. Thus, the major challenge in BrCa therapy is to improve the value of chemotherapy by reducing ADR resistance and identifying novel targets to reverse that resistance.

In recent years, significant breakthroughs have been made in BrCa therapy, particularly with the improvement of targeted therapies in the mechanisms of therapy resistance. Significantly, the effectiveness of endocrine therapy and patient prognosis have been improved by the use of the inhibitors of cyclin-dependent kinase 4 (CDK4) and cyclin-dependent kinase 6 (CDK6) and the mammalian target of rapamycin [5]. Additionally, pertuzumab, trastuzumab, or docetaxel combined with chemotherapy offer improved prognosis value for patients with Human Epidermal Growth

Journal of BIOLOGICAL REGULATORS and Homeostatic Agents

Table 1. 207 DEGs identified in GSE24460.

Expression level	DEGs (ranked in descending order of fold change)
Upregulated genes	*MMP1, VIM, ABCB1, LDHB, ABCB4, NNMT, FSTL1, CTGF, GPX1, UCHL1, AKAP12, PLS3, CYR61, CNN3, DKK3, CALD1, SPARC, MSN, SNURF, NEFH, SNRPN, AKR1B1, BEX1, TPM2, AXL, SERPINE1, PRSS21, IGF2BP2, FOXG1, EMP1, MCAM, COL1A1, PLAC8, PDGFC, COL4A2, LAMB1, GJA1, SRPX, RAI14, MYL9, FXYD5, NNMT, TCEAL9, IL32, SERPINE2, PALM2,NDN, RCN1, CDKN2A, XAGE1B, RAC2, GSTP1, TM4SF1, SLC12A8, AKAP2, KRT7, FBN2, PRAME, CRIP1, TFPI2, PTRF, ABC, GLYR1, IGF2BP3, IGFBP3, TGM2, CX-ADR, TGFB1I1, CDA, IL7R, PLAGL1, EXT1, HEG1, LOXL2, FYN, TGFBI, COL6A1, P3H2, SDC2, SLC16A3, EMP3, FOSL1, IFI16, ICAM1, NT5E, CTSZ, SLC16A4, PSG9, ANXA1, DFNA5, BCAT1, TUBB6, GNAI1, FABP5, CDH2, FKBP1B, FHL2, COL6A2, IGFBP, PEG10, SNCA, AKT3, COL4A1, CRYAB, BDNF, ZBTB18, MLLT11, ADGRL2, CLGN, STXBP6, MSLN*
Downregulated genes	*AGR2, UBB, TFF1, CDH1, KRT19, B2M, S100P, ESR1, GLUL, SLC6A14, LPCAT1, PDCD10, FXYD3, GREB1, DNAJC15, CA2, SCUBE2, MGP, MYO5C, CEACAM5, ID2, KYNU, LXN, MSMB, TMEM30B, B3GALNT1, CEA-CAM6, GALNT3, GFRA1, GATA3, NEBL, ESRP1, S100A14, ABAT, MYB, NRCAM, TACSTD2, KLF4, SH3YL1, CALCR, RB1, MAP7, CD24, HIST1H2AC, FBP1, GNA14, CYP1B1, AREG, RLN2, NCOA3, BMP7, OSR2, FOXA1, SLC24A3, CXCL12, ADCY1, SLC27A6, SYTL2, SCD, CELSR2, MREG, BRIP1, SLC2A10, IRS1, GDF15, HIST2H2AA4, ANK3, EPCAM, WWOX, HIST1H2BD, IGFBP5, PPP1R3D, MPPED2, SELENBP1, UGT2B15, RAB25, HNMT, HIST2H2BE, PBX1, XBP1, ST6GALNAC2, ISOC1, CLDN3, MSX2, SLC39A6, UGDH, MAFB, INSIG1, GSE1, RBPMS, TTC39A, TFF3, MTUS1, SERPINA3, RBM47, ANXA9*

Factor Receptor 2 (HER2)-positive status [6].

However, no ADR-resistant prognostic biomarkers are available for widespread clinical use, and the mechanisms of acquired resistance to ADR remain unknown. As a result, the goal of this research was to identify the major predicted pathways and genes as biomarkers/targets in the therapy of patients with ADR resistance in order to improve their prognosis.

Using microarray gene expression profiling, there is association of pathways of BrCa of the resistance acquired to ADR with the differentially expressed genes (DEGs) and pathways. To find the predictive genes in ADR resistance, functional enrichment, pathway, and PPI analyses were also carried out on the DEGs. Additionally, these crucial DEGs were examined in terms of relapse-free survival and the overall survival (OS) using clinical data from BrCa patients (RFS). To emphasize the differences even more, the constant critical genes in MCF-7 (cell lines in the human breast cancer) and MCF-7/ADR cells and tissues were discovered. The identification of these genes may help to better understand the mechanism of clinically observed chemotherapy resistance and could help identify BrCa cases in which chemotherapy has no effect on response or survival.

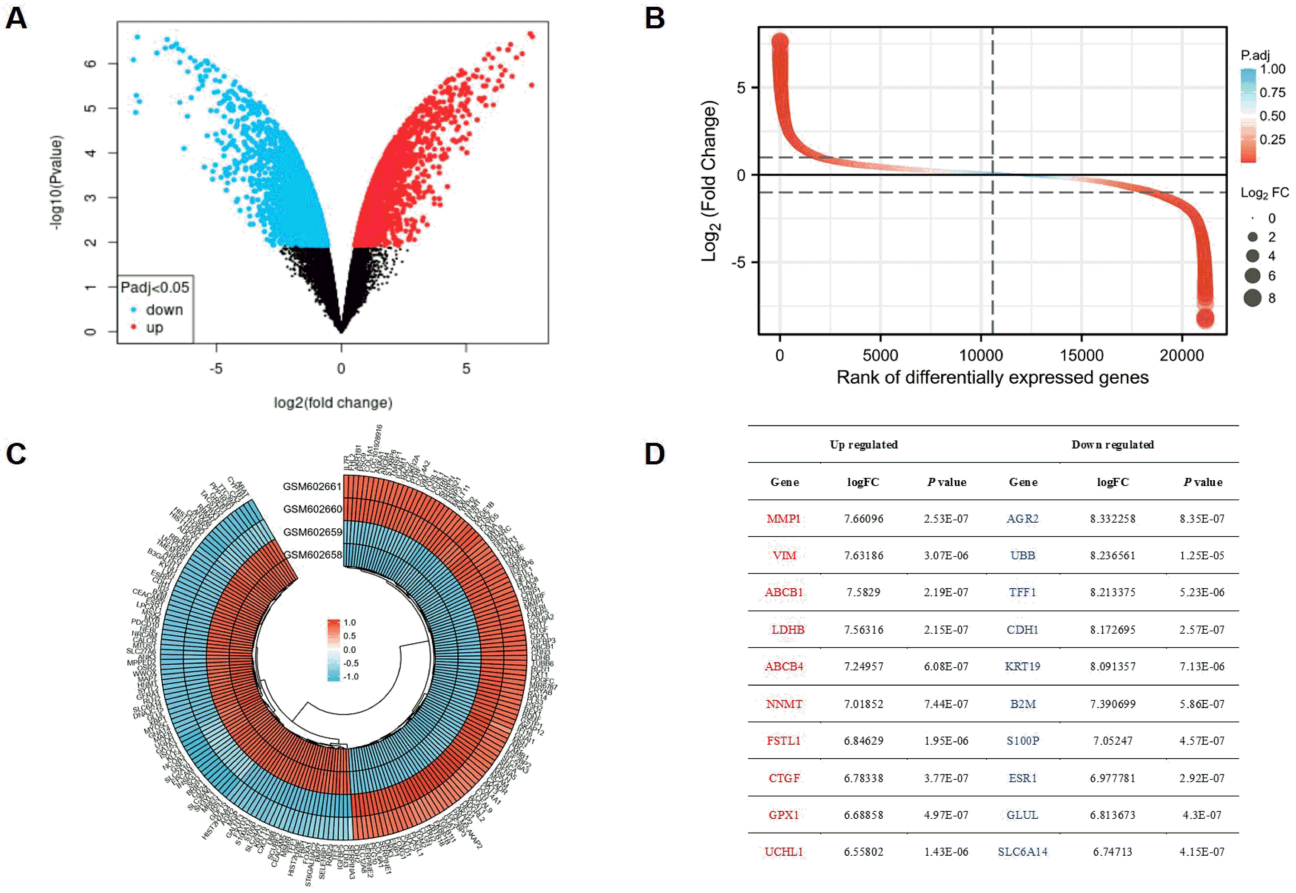
Results

Analysis of Differentially Expressed Genes in Adriamycin-Resistant Breast Cancer Using GEO2R

We used data (22,277 probe sets) derived from the GSE24460 dataset on the GPL571 oligonucleotide microarray platform in conjunction with the GEO2R tool in the database of the Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/). A total of 207 DEGs (111 upregulated and 96 downregulated in the ADR-resistant group) were notably found to be related to ADR resistance ($p < 0.05$ and $|\log_2 \text{Fold Change (FC)}| \geq 4$) (Table 1). Then, we extracted gene expression values and compiled a volcano plot (Fig. 1A) and a difference sort map (Fig. 1B) to show the distribution and expression of all the DEGs. The DEGs with $|\log_2 FC| \geq 4$ and $p < 0.05$ were presented in a ring-shaped heat map (Fig. 1C). Matrix metallopeptidase 1 (*MMP1*), ATP binding cassette subfamily B member 1 (*ABCB1*), vimentin (*VIM*), lactate dehydrogenase B (*LDHB*), ATP binding cassette subfamily B member 4 (*ABCB4*), nicotinamide N-methyltransferase (*NNMT*), follistatin like 1 (*FSTL1*), and connective tissue growth factor (*CTGF*) were the top ten upregulated (*UCHL1*). Priority gradient 2 (*AGR2*), ubiquitin B (*UBB*), trefoil factor 1 (*TFF1*), cadherin 1 (*CDH1*), keratin 19 (*KRT19*), beta-2-microglobulin (*B2M*), S100 calcium binding protein P (*S100P*), estrogen receptor 1 (*ESR1*), glutamate-ammonia ligase (*GLUL*), and solute carrier family 6 member 14 (*SLC6A14*) were the top (Fig. 1D).

Function and Pathway Enrichment Analysis

We uploaded the analysis of the Gene Ontology (GO) on the 207 DEGs, the Database for Annotation, Integrated Discovery (DAVID) and visualization. As shown in Fig. 2A, this was categorized into three parts: Cellular components (CCs), molecular functions (MFs), and biological processes (BPs). The responses of the DEGs to antibiotics, the control of cell-cell adhesion, reactive oxygen species, and hydrogen peroxide were the main areas where the BPs of the DEGs were enriched. There was a significant enrichment in the extracellular matrix (ECM), endoplasmic reticulum lumen, focal adhesion, lateral plasma membrane, and other CCs of DEGs. When it came to MF, DEGs were enriched for ECM structural components, cell adhesion me-



Up regulated			Down regulated		
Gene	logFC	*P* value	Gene	logFC	*P* value
MMP1	7.66096	2.53E-07	AGR2	8.332258	8.35E-07
VIM	7.63186	3.07E-06	UBB	8.236561	1.25E-05
ABCB1	7.5829	2.19E-07	TFF1	8.213375	5.23E-06
LDHB	7.56316	2.15E-07	CDH1	8.172695	2.57E-07
ABCB4	7.24957	6.08E-07	KRT19	8.091357	7.13E-06
NNMT	7.01852	7.44E-07	B2M	7.390699	5.86E-07
FSTL1	6.84629	1.95E-06	S100P	7.05247	4.57E-07
CTGF	6.78338	3.77E-07	ESR1	6.977781	2.92E-07
GPX1	6.68858	4.97E-07	GLUL	6.813673	4.3E-07
UCHL1	6.55802	1.43E-06	SLC6A14	6.74713	4.15E-07

Fig. 1. Identification of differentially expressed genes in MCF-7/ADR and MCF-7 breast cancer cells. (A) GSE24460 database volcano map of differentially expressed genes (DEGs). The X-axis is $\log_2$FC, while the Y-axis is log10 (*p* value). Blue: Downregulated genes, red: Upregulated genes, black: Non-significant genes. (B) Difference sort map of DEGs in GSE24460. X-axis: Rank of DEGs, Y-axis: $\log_2$FC. The closer to the left and right sides, the greater the absolute value of the difference. (C) DEG ring-shaped heat map in GSE24460 with $|\log_2 FC| \geq 4$ and $p < 0.05$. The white graphs show no difference in expression, the blue plots show downregulated genes, and the red plots show upregulated genes. (D) The top ten most significantly upregulated and downregulated genes in MCF-7/ADR cells.

diator activity, the activity of the cell-cell adhesion mediator, cadherin binding implicated in cell-cell adhesion, and so on.

The Kyoto Encyclopedia on the pathways of Genes and Genomes (KEGG) reference and the DAVID online tool were also used to classify the 207 DEGs. The analysis of A KEGG revealed that the 207 DEGs were enhanced in signaling pathways for focal adhesion, relaxin and estrogen signaling, leukocyte transendothelial migration, and the developed glycosylation receptor as the end product for advanced glycosylation end-product that shows pathway in diabetics (Fig. 2B). Fig. 2C depicts the top ten most significantly enriched GO keywords as well as the clustering of DEGs based on the *p* value. Fig. 2D shows the KEGG chord plot showing the clustering of the DEGs and their 8 corresponding KEGG keywords. **Supplementary Tables 1,2** summarize and display the top ten enriched GO keywords and pathways. The following 12 flapping DEGs may be linked with ADR resistance are found in the 10 GO-enriched words and 8 KEGG pathways: Collagen type I alpha 1 chain (COL1A1) MMP1, retinoblastoma (RB1), C-X-C motif ligand 12 (CXCL12), intercellular cell adhesion molecule-1 (ICAM1), collagen type VI alpha 1 chain (COL6A1), insulin receptor substrate 1 (IRS1), collagen type I alpha 1 chain (COL1A1), and cyclin-dependent kinase (CDKN2A).

Analysis of Protein–Protein Interaction Networks

For the data of PPI to be obtained, we did through retrieving of the Interacting Genes/Protein from the website through the use search tool. Next, samples with PPI data values of >0.9 were selected to assemble PPI networks using Cytoscape software v. 3.4.0 (National Institute of General Medical Sciences; Bethesda, MD, USA). The STRING tool extracted 393 PPI pairs containing 150 DEGs, including 74 upregulated genes and 76 downregu-

Journal of
BIOLOGICAL REGULATORS
and Homeostatic Agents

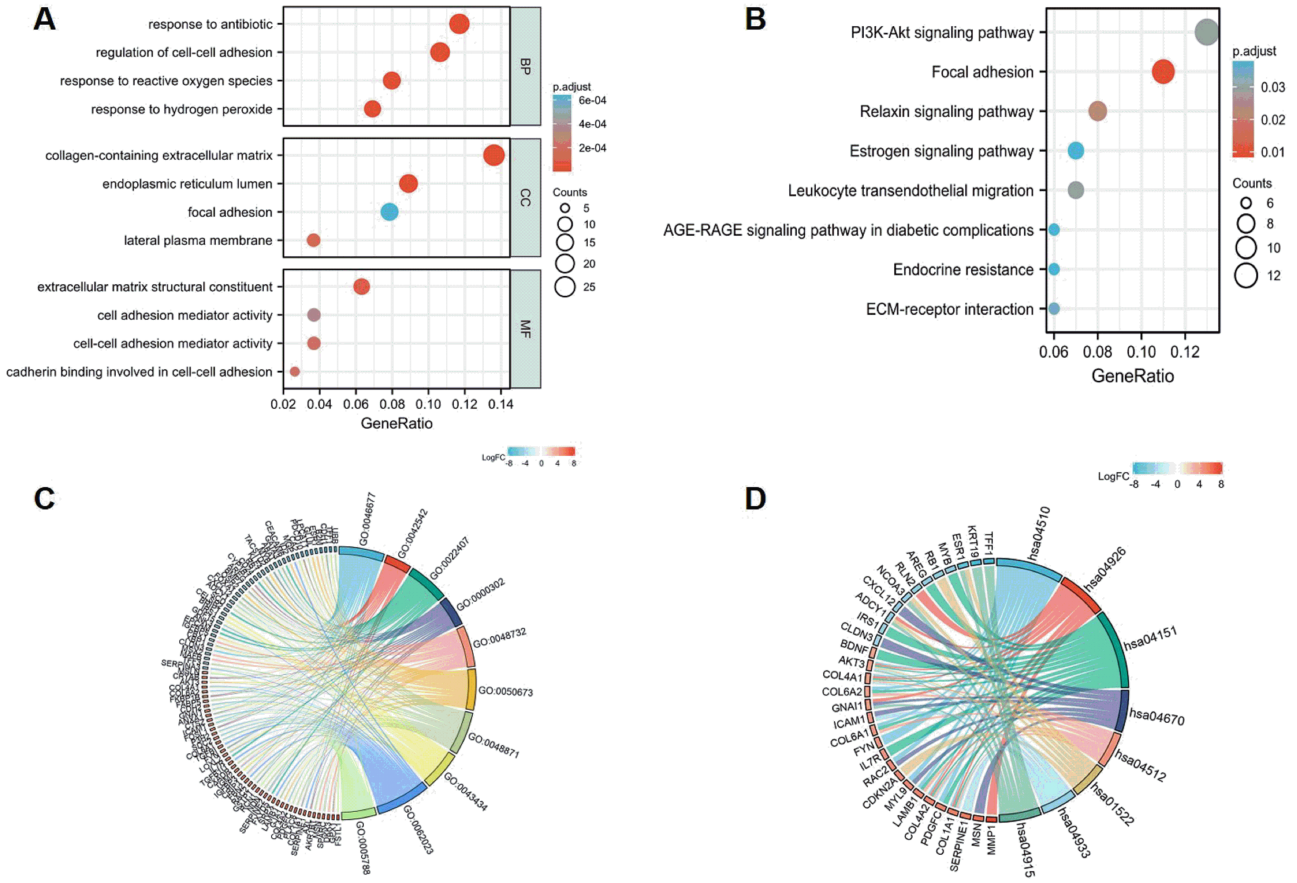


Fig. 2. Gene Ontology enrichment and Kyoto Encyclopedia of Genes and Genomes pathway analysis of differentially expressed genes in MCF-7/Adriamycin-resistant and MCF-7 breast cancer cells. (A) Gene Ontology (GO) annotation diagram of 207 DEGs with $|\log_2 FC| \geq 4$ and $p < 0.05$, comprising the GO categories of biological processes, cellular components, and molecular activities. (B) Pathway annotation diagram of 207 DEGs from the Kyoto Encyclopedia of Genes and Genomes (KEGG). (C) A GO chord plot of DEG clustering and the assigned top 10 GO terms. The DEGs are linked to their assigned terms by colored ribbons. The blue–red coding next to the selected genes represents $\log_2 FC$. (D) A KEGG chord plot of DEG clustering and the assigned top 8 KEGG terms. The DEGs are linked to their assigned terms by colored ribbons. The blue–red coding next to the selected genes represents $\log_2 FC$.

lated genes. To assess their linkages, these genes were normally distributed in a single PPI picture (Fig. 3A). Furthermore, when the cut-off criterion was set to "Degrees 10", the following 10 genes in the network of PPI were identified as hub genes: *VIM, CDH1, ESR1, CDKN2A, CXCL12, RB1, UBB, ICAM1, COL4A1*, and serpin family E member 1 (*SERPINE1*).

Genes Correlated with Adriamycin Resistance in Patients with Breast Cancer and Corresponding Survival Analysis

The GO enrichment analysis genes, KEGG-enriched genes, and PPI hub genes yielded a total of 6 overlapping genes (Fig. 4A). We used box plots to analyze gene expression between a neoadjuvant chemotherapy-sensitive group and a resistant group in the GSE34138 dataset from the GEO to determine the functions of the overlapping genes in ADR resistance. The expression of the six most overlapped genes was generated from the GSE34138 RNA-sequence dataset, which comprised 178 patients treated with neoadjuvant chemotherapy. Among the groups that is sensitive and the resistant of the two genes exhibited significant differences; In the latter, *COL4A1* was increased, while *CDH1* was decreased (Fig. 4B).

Next, the method used to analyze the medical value of the two genes is Kaplan–Meier (KM). As the data in Fig. 4C revealed, the lower expression of *COL4A1* predicted longer OS and RFS times in the patients with BrCa ($p = 0.014$, $p = 5.4 \times 10^{-5}$), and *CDH1* expression was shown to be substantially related to RFS ($p = 3.5 \times 10^{-5}$).

COL4A1 Expression and Impact on Sensitivity to Adriamycin in MCF-7/Adriamycin Cells

Then, we discovered that *COL4A1*, which is highly expressed in BrCa tissue that is drug-resistant, was expressed in MCF-7 and MCF-7/ADR cells. Using the reverse transcription polymerase chain reaction (RT-PCR)

A

B

Gene name	Degree	Betweeness
VIM	34	0.28575217
CDH1	34	0.20156051
ESR1	32	0.19746661
CDKN2A	15	0.05243322
RB1	14	0.05094056
CXCL12	14	0.04287146
UBB	13	0.0501096
ICAM1	13	0.03445037
SERPINE1	12	0.07012124
COL4A1	12	0.0481231

Fig. 3. Protein–protein interaction network and hub genes. (A) STRING database protein-protein interaction network of differentially expressed genes (DEGs). Upregulated gene expression is represented by red plots, whereas downregulated gene expression is represented by green plots. DEG interactions are represented by lines. (B) The hub genes discovered and their interactions.

Journal of
BIOLOGICAL REGULATORS
and Homeostatic Agents

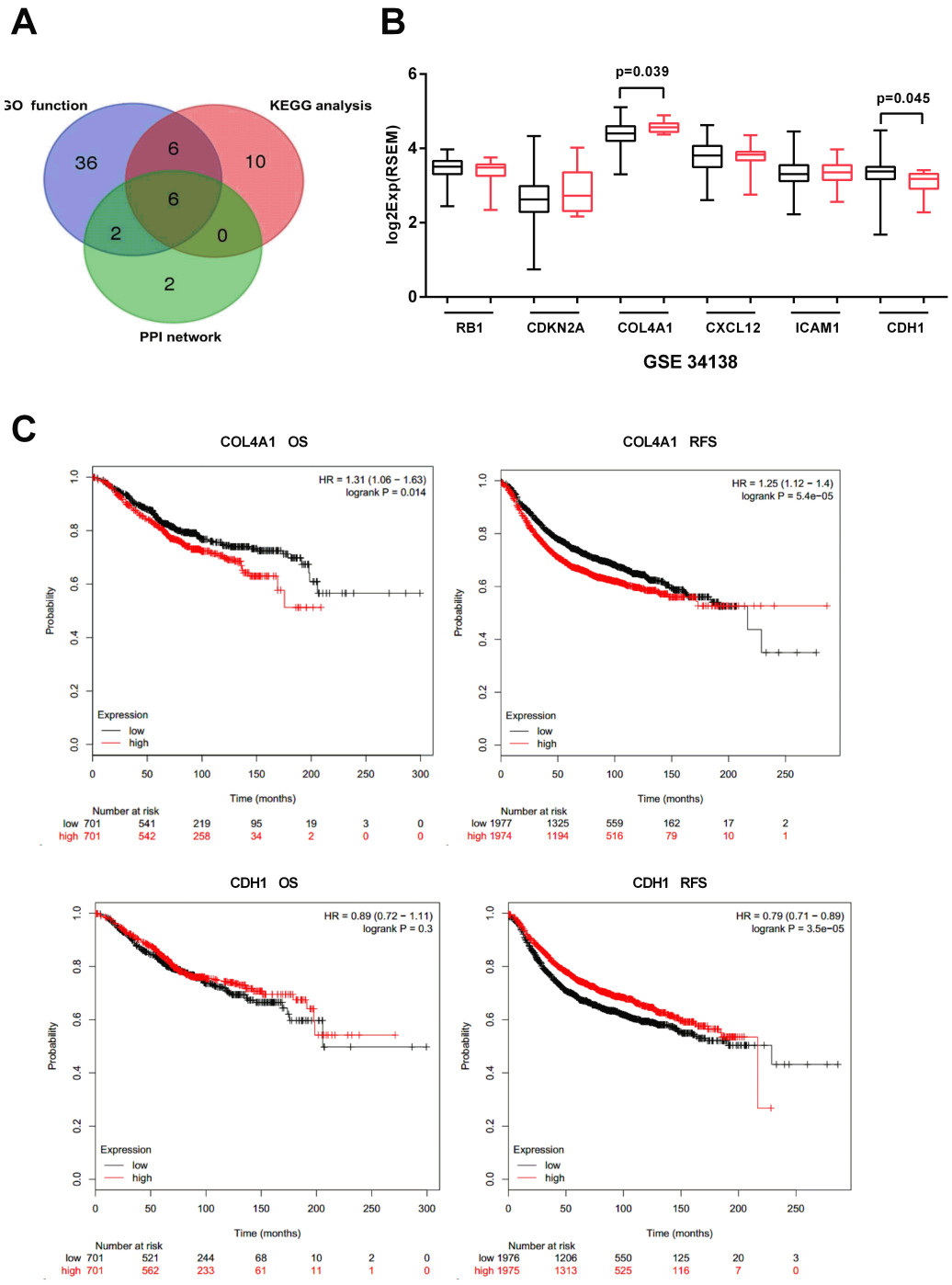


Fig. 4. Genes correlated to patients with Adriamycin-resistant breast cancer and a survival analysis. (A) Venn diagram of overlapping differentially expressed genes (DEGs) identified using Gene Ontology enrichment studies, Kyoto Encyclopedia of Genes and Genomes pathway analyses, and protein-protein interactions. (B) Box plots displaying gene expression of six candidate genes (RB transcriptional corepressor 1 (*RB1*), cyclin dependent kinase inhibitor 2A (*CDKN2A*), C-X-C motif chemokine ligand 12 (*CXCL12*), collagen type IV alpha 1 chain (*COL4A1*), intercellular adhesion molecule 1 (*ICAM1*), and cadherin 1 (*CDH1*) in patients with anthracycline-based (BrCa)). The black lines and red lines represent patients with chemotherapy sensitivity and chemotherapy resistance, respectively. The RNA-sequence dataset GSE34138 was used. The patients (n = 178) were separated into two groups: Sensitive and resistant, and $p < 0.05$ was considered significant. (C) Kaplan–Meier (KM) survival curves comparing the predictive value of *CDH1* and *COL4A1* expression levels in Adriamycin resistance in terms of overall survival and relapse-free survival. The curves were created using an online application called KM plotter. The levels of DEG expression were dichotomized using the median value, and the findings were visualized using KM survival plots. Log-rank statistics were used to compute the *p* values (Number of patients = 1402).

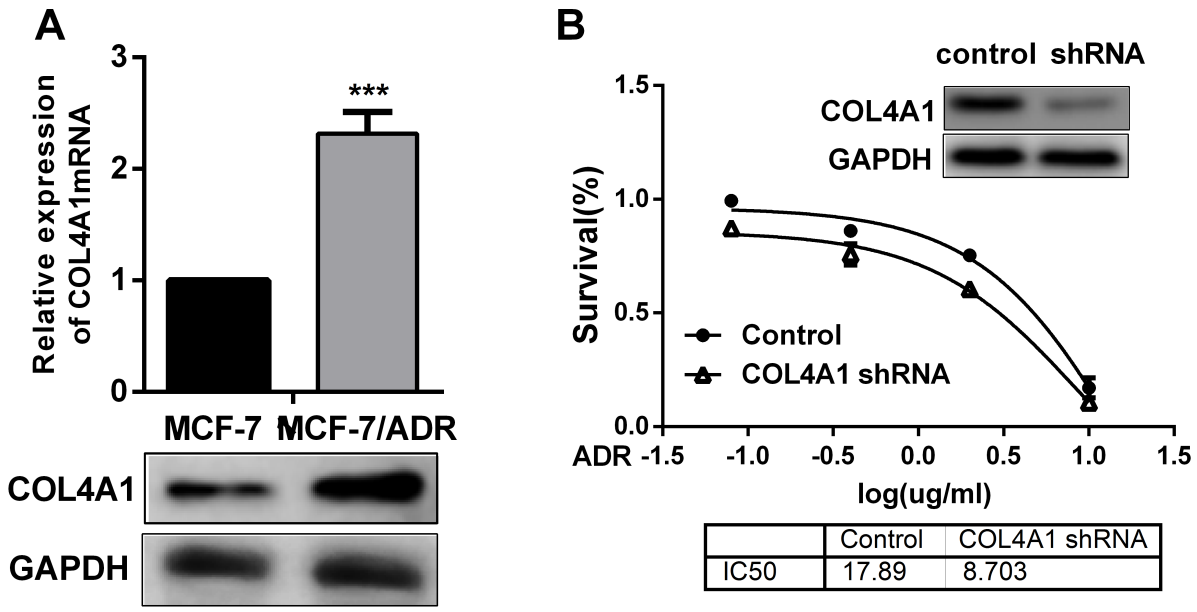


Fig. 5. ***COL4A1*** **expression and impact on sensitivity to Adriamycin in MCF-7/ADR cells.** (A) Quantitative real-time reverse-transcription polymerase chain reaction and Western blot analyses were used to assess *COL4A1* expression levels in MCF-7 and MCF-7/Adriamycin (ADR) cells. As a loading control, glyceraldehyde-3-phosphate dehydrogenase was used ($^{***}p < 0.001$). (B) After 48 hours, MCF-7/ADR cells were transfected with *COL4A1* shRNA and treated with several concentrations of ADR (0.00, 0.08, 0.40, 2.00, and 10.00 μg/mL). A Cell Counting Kit-8 assay was used to count the surviving cells ($^{*}p < 0.05$).

method, it was discovered that the higher level of *COL4A1* was found MCF-7/ADR in messenger RNA expression than the cells in MCF-7. The confirmation of the results was done by Western blot analysis (Fig. 5A). The GEO analysis was consistent with the differences in *COL4A1* expression between the Western blot data and RT-PCR.

As a result of strong correlation between *COL4A1* and OS/RFS, we were able to determine that *COL4A1* was involved in the growth of MCF-7/ADR cells. In MCF-7/ADR cells, *COL4A1* knockdown significantly reduced the concentration of inhibitory to 50% (IC50) value of ADR, whereas *COL4A1* overexpression significantly increased the IC50 of ADR (**Supplementary Fig. 1**). Therefore, it was clear from our research that *COL4A1* controlled BrCa cell growth and was crucial for ADR resistance. Therefore, *COL4A1* may be a new target that can be promising for enhancing the patients' prognosis with BrCa that is resistant to ADR.

Discussion

The most dangerous women tumor in the world is breast cancer, posing a major morbidity and mortality risk to patients' health and lives [7]. Its main treatment strategy is chemotherapy, and usually, ADR is the major agent for patients with BrCa, especially in advanced phases [8]. However, the major obstance to the treatment of BrCa is the resistance that is developed from the ADR and leads to cancer-related deaths [9]. The mechanisms giving rise to the progression of ADR resistance are both complex and lack characteristics, and a better knowledge of the development of ADR resistance and its key genes will contribute to enhancing the therapeutic effect in patients with BrCa.

In our study, we examined the GSE24460 dataset's microarray data and discovered in the cells of 207 DEGs in MCF-7 and MCF-7/ADR. We hypothesized that these genes were likely linked to ADR resistance in BrCa patients. We used GO enrichment analysis, the analysis through KEGG pathway, and PPI network studies to examine the 207 DEGs. We discovered a clear link between DEGs and antibiotic response using GO enrichment and KEGG pathway analysis; Also, the majority of the genes were found in the focal adhesion pathway. Accordingly, we speculated that ADR resistance might be associated with interactions between BrCa cells.

In the GSE24460 and GSE34138 databases, two genes (*CDH1* and *COL4A1*) demonstrated significant differences between an ADR-sensitive group and an ADR-resistant group. These are the new candidate genes for ADR resistance, in which *CDH1* was downregulated and *COL4A1* was upregulated in the ADR-resistant group. We further investigated the association of *CDH1* and *COL4A1* with OS and RFS and discovered that only *COL4A1* was correlated with both OS and RFS, while *CDH1* was associated only with RFS. A KM plotter analysis was conducted based on two groups of general patients with BrCa who were receiv-

Journal of
BIOLOGICAL REGULATORS
and Homeostatic Agents

ing or not receiving ADR therapy, which was the possible reason for the inconsistency. These analytical results should be verified further using the data of specific patients receiving ADR therapy.

Our findings show that *COL4A1* is related with both OS and RFS at the same time, putting it in a critical position. The *COL4A1* gene encodes a type-IV collagen alpha protein that is found in membranes of the basement and is conserved across species. This gene shares a bidirectional promoter with a homologous gene in another chain [10]. It produces a marked effect as part of a heterotrimer and interacts with other ECM compositions, including nacrose, proteoglycans, and laminins, to control both structural uniqueness and protein expression [11]. Herein, we discovered that *COL4A1* is significantly related to ADR resistance and is a poor prognostic marker for BrCa.

This result is consistent with earlier studies that discovered *COL4A1*, the identification of the people at risk for progression and recurrence of bladder cancer, papillary thyroid carcinoma, intrahepatic cholangiocarcinoma, and prostate cancerwas through the diagnostic and prognostic biomarker [12,13]. Furthermore, *COL4A1* has been linked to the proliferation, development, and advancement of BrCa [14]. Moreover, the mutation of *COL4A1* increased the risk of intracranial hemorrhage from disrupted integrity due to mutations in the vascular basement membrane and suspended myofibroblast proliferation and differentiation [15]. These results corresponded with our analytical findings in that *COL4A1* was upregulated in the ADR-resistance cohort and was used to predict the poor prognosis of BrCa patients.

Additionally, we discovered that *CDH1* is associated with BrCa ADR resistance and has a prognostic effect for RFS. Breast cancer progression or tumorigenesis is linked to mutations or losses in the *CDH1* gene, which encodes a classical cadherin of the cadherin superfamily. [16], gastric [17], ovarian [18], and colorectal [19] cancer. Additionally, it was reported that *CDH1* gene mutations could contribute to the development of hereditary diffuse gastric cancer, which affects the function of E-cadherin [20].

Material and Methods

Microarray Data

The Affymetrix Human Genome U133 Plus 2.0 Array GEO GPL571 platform provided the GSE24460 microarray dataset [21–23] (Affymetrix, Santa Clara, CA, USA). The GSE24460 dataset included (in duplicate) two parental human BrCa cell lines of MCF-7 and their acquired-resistance subset MCF-7/ADR, which was generated after exposure to high-dose ADR (860 nM) on every other passage. Then, using the SuperScriptTM system 3 (Invitrogen Life Technologies, Carlsbad, CA, USA), the cells were extracted from the total RNA and converted to cDNA using a T7-(dT)24 primer.

Data Processing and Screening of Differentially Expressed Genes

In comparison of two or more groups of samples, the GEO2R online application (https://www.ncbi.nlm.nih.gov/geo/geo2r/) was used. To detect DEGs between experimental groups, gene expression comparisons were performed in the GEO dataset [24]. In this study, DEGs were found by contrasting their expression in MCF-7 parental cell lines with related ADR-resistant cell lines. The Benjamini-Hochberg procedure was used to correct the microarray's false positive result, and modified *p* values were used to calculate the *p* values [25]. The terms "adj. *p* value < 0.05" and "$|\log_2 FC| \geq 4$" were used as check standards.

Enrichment Analysis of Differentially Expressed Genes

The DAVID database (https://david.ncifcrf.gov) [26] is an online database that can be used to identify significant DEGs by their BPs, MFs, or CPs using GO comprehensive categorical data (http://www.geneontology.org/) [27] as well as identify important transcripts using DAVID's functional annotation clustering tool (Benjamini-Hochberg FDR 0.05). Furthermore, the DAVID database was used for pathway enrichment analysis in conjunction with the KEGG database (http://www.genome.jp/kegg/), with a Benjamini-Hochberg FDR of 0.05 as the cut-off threshold [28].

Construction of Protein–Protein Interaction Networks

The STRING website (http://www.string-db.org/) was used to build the PPI networks [29] to forecast protein interaction connections. We chose a composite score of >0.9 (high confidence) as the cut-off standard in this study. The hub genes were chosen based on the node degree, and the PPI network was visualized using Cytoscape software v. 3.4.0 [30] (http://www.cytoscape.org) (National Institute of General Medical Sciences; Bethesda, MD, USA). The number (degree) of lines connecting proteins in Cytoscape was evaluated to identify hub proteins with essential biological roles, as well as the number of nodes that were not directly connected by a given node (i.e., the betweenness value) for each node.

Confirmation of Adriamycin-Resistant Genes

Gene expression profiles were acquired from the GSE34138 [31] RNA-sequence dataset to study the involvement of important genes in predicting ADR resistance (GSE24460, U133 Plus 2.0, Affymetrix, Santa Clara, CA, USA). Patients undergoing anthracycline-based chemotherapy were separated into two groups: Those who responded to the treatment were labeled chemotherapy sensitive, while those who did not respond were labeled chemotherapy resistant. We analyzed 178 breast tumor biopsies from patients in the GSE34138 dataset, in which

the gene expression value was shown in a $\log_2$-transformed format. The expression of the key genes was compared between the patients with chemotherapy sensitivity and those with chemotherapy resistance, and the significance value was calculated using a Mann–Whitney U test an esteem of $p < 0.05$ was measurably considered critical.

Cell Culture and Transfection

The American Type Culture Collection provided the human BrCa MCF-7 cell line and ADR-resistant MCF-7 cells MCF-7/ADR (MCF-7/ADR). In Dulbecco's Modified Eagle Medium (Invitrogen, Carlsbad, CA, USA), cells were maintained, 10 μL bovine serum (HyClone, Logan, UT, USA), 100 U/mL penicillin and 0.1 mg/mL streptomycin in 37 °C in a humidified container. environment with 5% CO_2. In order to keep the MCF-7/ADR cells' ADR-resistant phenotype growth medium, 100 nM ADR was added. The cells of MCF-7/ADR were transfected with 50-nM *COL4A1* short-hairpin (sh) RNA and the negative control using Lipofectamine 2000 (scrambled control oligonucleotides). MCF-7 cells were transfected with 4 g of *COL4A1* cDNA and the negative control using Lipofectamine 2000 (scrambled control oligonucleotides). For the Western blot analysis and quantitative RT-PCR, cells were collected in 48 hours after transfection.

Quantitative Real-Time Reverse-Transcription Polymerase Chain Reaction and Western Blot

Total RNA was collected by the use TRIZOl reagents (Invitrogen, Carlsbad, CA, USA) from MCF-7 and MCF-7/ADR cells under cultivation in accordance with the manufacturer's instructions. Utilizing a SYBR® Green-based PCR Master Mix, qRT-PCR was carried out (Life Technologies, Carlsbad, CA, USA). In **Supplementary Table 3**, the primers are listed. *COL4A1* expressed itself relative to GAPDH. The method of 2-Ct was used to determine the fold change for each RNA in comparison to the control.

Total protein was separated using a proteinase-supplemented RIPA (Radio Immunoprecipitation Assay) buffer (1-mM EthyleneDiamineTetraacetic Acid (EDTA), 150-mM NaCl, 50-mM Tris-HCl at pH 7.4, 1% NP-40, and 0.5% Sodium dodecyl sulfate (SDS)) (Roche, Basel, Switzerland). The solution was heated to 100 °C for five minutes. Using 10% SDS polyacrylamide gels, the denatured protein samples were loaded, separated, and transferred to a polyvinylidene fluoride membrane. After 1 hour of blocking with 5% skimmed milk powder in Tris-buffered saline/0.05% Tween 20, the membrane was incubated overnight at 4 °C with a specific primary antibody. After that, secondary antibodies were applied to the membrane and left on it at room temperature for an hour. Using the antibodies in **Supplementary Table 4**, the Supplementary was described. To see the bands, the chemiluminescent detecting apparatus was employed (AI680, Amersham, Freiburg, Germany). The analysis of the western blot bands' intensity was conducted using ImageJ software (version 1.48, National Institutes of Health, Bethesda, MA, USA) (http://rsb.info.nih.gov/ij).

Cell Proliferation

The 50-nM *COL4A1* shRNA and the negative control were transfected into MCF-7/ADR cells in accordance with the manufacturer's instructions. 4 g of *COL4A1* cDNA and the negative control were transfected into MCF-7 cells using Lipofectamine 2000 as directed by the manufacturer. ADR was introduced to the cells at a range of concentrations (0.00, 0.08, 0.40, 2.00, and 10.00 g/mL) after transfection for 48 hours. The absorbance at 450 nm was measured using a Cell Counting Kit-8 test, and the number of remaining cells was counted (Dojindo, Japan).

Conclusions

The above findings suggest new target genes for predicting the emergence of ADR resistance. Moreover, the enriched target genes could also be applied to predict the prognosis of patients with BrCa. However, whether these target genes can be used in clinical therapy depends on the validation of our results in future research.

Author Contributions

SL—Conception and design of the research; TL, SL—Acquisition of data; YX, QW, SL—Analysis and interpretation of the data; YX, PPG—Statistical analysis; SL—Obtaining financing; TL, SL, YX—Writing of the manuscript; PPG, QW—Critical revision of the manuscript for intellectual content. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

We are particularly grateful to all the people who have given us help on our article.

Funding

Natural Science Foundation of Liaoning Province Science and Technology Department (No. 2019-BS-236), Scientific Research Foundation of Liaoning Province Education Department (No. LJKZ1141), Scientific Research Foundation of Shenyang Medical College (No. 20195070), and Undergraduate Scientific Research Project of Shenyang Medical College (No. 20219019).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.23812/j.biol.regul.homeost.agents.20233704.223.

References

[1] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin.* 2022;72(1):7–33. doi: 10.3322/caac.21708

[2] Barzaman K, Karami J, Zarei Z, *et al*. Breast cancer: Biology, biomarkers, and treatments. *Int Immunopharmacol.* 2020;84:106535. doi: 10.1016/j.intimp.2020.106535

[3] Gupta GK, Collier AL, Lee D, *et al*. Perspectives on Triple-Negative Breast Cancer: Current Treatment Strategies, Unmet Needs, and Potential Targets for Future Therapies. *Cancers (Basel).* 2020;12(9):2392. doi: 10.3390/cancers12092392

[4] Austreid E, Lonning PE, Eikesdal HP. The emergence of targeted drugs in breast cancer to prevent resistance to endocrine treatment and chemotherapy. *Expert Opin Pharmacother.* 2014;15(5):681–700. doi: 10.1517/14656566.2014.885952

[5] Xi J, Ma CX. Sequencing Endocrine Therapy for Metastatic Breast Cancer: What Do We Do After Disease Progression on a CDK4/6 Inhibitor? *Curr Oncol Rep.* 2020;22(6):57. doi: 10.1007/s11912-020-00917-8

[6] Fleeman N, Bagust A, Beale S, *et al*. Pertuzumab in combination with trastuzumab and docetaxel for the treatment of HER2-positive metastatic or locally recurrent unresectable breast cancer. *Pharmacoeconomics.* 2015;33(1):13–23. doi: 10.1007/s40273-014-0206-2

[7] Ahmad A. Breast Cancer Statistics: Recent Trends. *Adv Exp Med Biol.* 2019;1152:1–7. doi: 10.1007/978-3-030-20301-6_1

[8] Leon-Ferre RA, Hieken TJ, Boughey JC. The Landmark Series: Neoadjuvant Chemotherapy for Triple-Negative and HER2-Positive Breast Cancer. *Ann Surg Oncol.* 2021;28(4):2111–2119. doi: 10.1245/s10434-020-09480-9

[9] Nedeljković M, Damjanović A. Mechanisms of Chemotherapy Resistance in Triple-Negative Breast Cancer-How We Can Rise to the Challenge. *Cells.* 2019;8(9):957. doi: 10.3390/cells8090957

[10] Kuo DS, Labelle-Dumais C, Gould DB. COL4A1 and COL4A2 mutations and disease: insights into pathogenic mechanisms and potential therapeutic targets. *Hum Mol Genet.* 2012;21(R1):R97–R110. doi: 10.1093/hmg/dds346

[11] Jeanne M, Gould DB. Genotype-phenotype correlations in pathology caused by collagen type IV alpha 1 and 2 mutations. *Matrix Biol.* 2017;57–58:29–44. doi: 10.1016/j.matbio.2016.10.003

[12] Miyake M, Morizawa Y, Hori S, *et al*. Diagnostic and prognostic role of urinary collagens in primary human bladder cancer. *Cancer Sci.* 2017;108(11):2221–2228. doi: 10.1111/cas.13384

[13] Steele R, Mott JL, Ray RB. MBP-1 upregulates miR-29b that represses Mcl-1, collagens, and matrix-metalloproteinase-2 in prostate cancer cells. *Genes Cancer.* 2010;1(4):381–387. doi: 10.1177/1947601910371978

[14] Jin R, Shen J, Zhang T, *et al*. The highly expressed COL4A1 genes contributes to the proliferation and migration of the invasive ductal carcinomas. *Oncotarget.* 2017;8(35):58172–58183. doi: 10.18632/oncotarget.17345

[15] Hino-Fukuyo N, Kikuchi A, Iwasaki M, *et al*. Dramatic response after functional hemispherectomy in a patient with epileptic encephalopathy carrying a de novo COL4A1 mutation. *Brain Dev.* 2017;39(4):337–340. doi: 10.1016/j.braindev.2016.11.006

[16] Girardi A, Magnoni F, Vicini E, *et al*. CDH1 germline mutations in families with hereditary lobular breast cancer. *Eur J Cancer Prev.* 2022;31(3):274–278. doi: 10.1097/CEJ.0000000000000688

[17] Shenoy S. CDH1 (E-Cadherin) Mutation and Gastric Cancer: Genetics, Molecular Mechanisms and Guidelines for Management. *Cancer Manag Res.* 2019;11:10477–10486. doi: 10.2147/CMAR.S208818

[18] Graffeo R, Livraghi L, Pagani O, Goldhirsch A, Partridge AH, Garber JE. Time to incorporate germline multigene panel testing into breast and ovarian cancer patient care. *Breast Cancer Res Treat.* 2016;160(3):393–410. doi: 10.1007/s10549-016-4003-9

[19] Ye CC, Wang J. E-cadherin (CDH1) gene -160C/A polymorphism and the risk of colorectal cancer: A meta-analysis involving 17,291 subjects. *J Gene Med.* 2021;23(10):e3370. doi: 10.1002/jgm.3370

[20] Nemtsova MV, Bykov II, Udilova AA, Zaletaev DV, Khorobrykh TV. [Allelic imbalance of loci 17p13.1 (TP53), 1p36.1 (RUNX3), 16p22 (CDH1) and microsatellite instability in gastric cancer]. *Mol Biol (Mosk).* 2013;47(5):835–841.

[21] Calcagno AM, Salcido CD, Gillet JP, *et al*. Prolonged drug selection of breast cancer cells and enrichment of cancer stem cell characteristics. *J Natl Cancer Inst.* 2010;102(21):1637–1652. doi: 10.1093/jnci/djq361

[22] Koch M, Wiese M. Quality Visualization of Microarray Datasets Using Circos. *Microarrays (Basel).* 2012;1(2):84–94. doi: 10.3390/microarrays1020084

[23] Ritchie ME, Phipson B, Wu D, *et al*. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47. doi: 10.1093/nar/gkv007

[24] Barrett T, Wilhite SE, Ledoux P, *et al*. NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res.* 2013;41(Database issue):D991–D995. doi: 10.1093/nar/gks1193

[25] Jary M, Hasanova R, Vienot A, *et al*. Molecular description of ANGPT2 associated colorectal carcinoma. *Int J Cancer.* 2020;147(7):2007–2018. doi: 10.1002/ijc.32993

[26] Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5):284–287. doi: 10.1089/omi.2011.0118

[27] Ashburner M, Ball CA, Blake JA, *et al*. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25(1):25–29. doi: 10.1038/75556

[28] Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 2016;44(D1):D457–D462. doi: 10.1093/nar/gkv1070

[29] Szklarczyk D, Franceschini A, Wyder S, *et al*. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 2015;43(Database issue):D447–D452. doi: 10.1093/nar/gku1003

[30] Shannon P, Markiel A, Ozier O, *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13(11):2498–2504. doi: 10.1101/gr.1239303

[31] de Ronde JJ, Lips EH, Mulder L, *et al*. SERPINA6, BEX1, AGTR1, SLC26A3, and LAPTM4B are markers of resistance to neoadjuvant chemotherapy in HER2-negative breast cancer. *Breast Cancer Res Treat.* 2013;137(1):213–223. doi: 10.1007/s10549-012-2340-x