

Bioinformatics-Driven Investigations of Signature Biomarkers for Triple-Negative Breast Cancer

Shristi Handa¹, Sanjeev Puri¹, Mary Chatterjee¹ and Veena Puri² 

¹Biotechnology Engineering, University Institute of Engineering and Technology, Panjab University, Chandigarh, India. ²Centre for Systems Biology and Bioinformatics, Panjab University, Chandigarh, India.

Bioinformatics and Biology Insights
Volume 19: 1–15
© The Author(s) 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/11779322241271565



ABSTRACT: Breast cancer is a highly heterogeneous disorder characterized by dysregulated expression of number of genes and their cascades. It is one of the most common types of cancer in women posing serious health concerns globally. Recent developments and discovery of specific prognostic biomarkers have enabled its application toward developing personalized therapies. The basic premise of this study was to investigate key signature genes and signaling pathways involved in triple-negative breast cancer using bioinformatics approach. Microarray data set GSE65194 from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus was used for identification of differentially expressed genes (DEGs) using R software. Gene ontology and Kyoto Encyclopedia of Genes and Genome (KEGG) pathway enrichment analyses were carried out using the ClueGO plugin in Cytoscape software. The up-regulated DEGs were primarily engaged in the regulation of cell cycle, overexpression of spindle assembly checkpoint, and so on, whereas down-regulated DEGs were employed in alteration to major signaling pathways and metabolic reprogramming. The hub genes were identified using cytoHubba from protein-protein interaction (PPI) network for top up-regulated and down-regulated DEGs plugin in Cytoscape software. The hub genes were validated as potential signature biomarkers by evaluating the overall survival percentage in breast cancer patients.

KEYWORDS: breast cancer, differentially expressed genes, Cytoscape, protein-protein interaction network, genes, ClueGo

RECEIVED: February 14, 2024. **ACCEPTED:** June 29, 2024.

TYPE: Research Article

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by, DBT Builder Grant BT/INF/22/SP41295/2020. The financial support provided by Panjab University is duly acknowledged. Infrastructure support by Department of Biotechnology, UIET and Centre for Systems Biology and Bioinformatics, Panjab University, Chandigarh, India is duly acknowledged.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Veena Puri, Centre for Systems Biology and Bioinformatics, Panjab University, Chandigarh 160014, India. Email: vpuri1825@pu.ac.in

Introduction

Breast cancer is one of the most commonly diagnosed cancers globally with about 2.26 million cases reported in 2020. Although this disease is considered to be largely predisposed in developed countries, but more than half of breast cancer cases were diagnosed and two third of breast cancer-related deaths were reported in the less developed regions.^{1,2} Of the known subtypes, triple-negative breast cancer (TNBC) is one of the most aggressive subtypes, accounting for 10% to 15% of all breast cancers. It is mainly characterized by the negative expression of estrogen (ER) and progesterone (PR) and human epidermal factor receptors (HER2).³ The biological and clinical characteristics endowed in TNBC's together with its aggressive nature disposes even younger women to experience deterioration and visceral metastases in comparison to other subtypes of breast cancer. Chemotherapy and radiation therapy remain the traditional therapeutic options for TNBC.⁴ Studies have shown that, despite advances in pharmacologic research and disease treatments, still there are no effective treatments available for TNBC. The situation is more concernful, as the chances of recurrence and mortality of the patients are still not effectively controlled. As a result, TNBC continues to be a major challenge due to its poor prognosis, high probability rate of metastasis and limited curative options.⁵

In recent years, integrated bioinformatics has been extensively applied in cancer research.⁶ With continuous advancements, the bioinformatics tools have now frequently been exploited in the diagnosis, prognosis, and screening of cancers.⁷ Numerous studies have shown the potential of integrating and using data sets to investigate the expression levels of several genes to predict prognosis of various cancer types.^{8,9} Thus, the use of bioinformatics in cancer research has opened up new avenues both for the development of diagnostic and prognostic markers as well as their use for improving the outcomes of cancer patients.^{10,11} A great progress has been achieved in understanding the genes that are differentially expressed in malignancies such as renal cell carcinoma,¹² prostate,¹³ ovarian,¹⁴ and other associated disorders.^{15–17} Many studies have found overexpression of CCNB1 in various cancers, such as pancreatic carcinoma¹⁸ and lung carcinoma.¹⁹ Likewise, Shi et al²⁰ demonstrated that ISL1-induced gastric cancer cell proliferation and carcinogenesis are mediated through the alteration in the expressions of genes like CCNB1, CCNB2, and C-MYC.

Breast cancer occurrence and advancement are a multistage process involving the synergistic action of number of genes. As a result, alteration in expression of these gene expressions has drawn interest in breast cancer for a long time.²¹ Currently, some research investigations have looked into the genes that



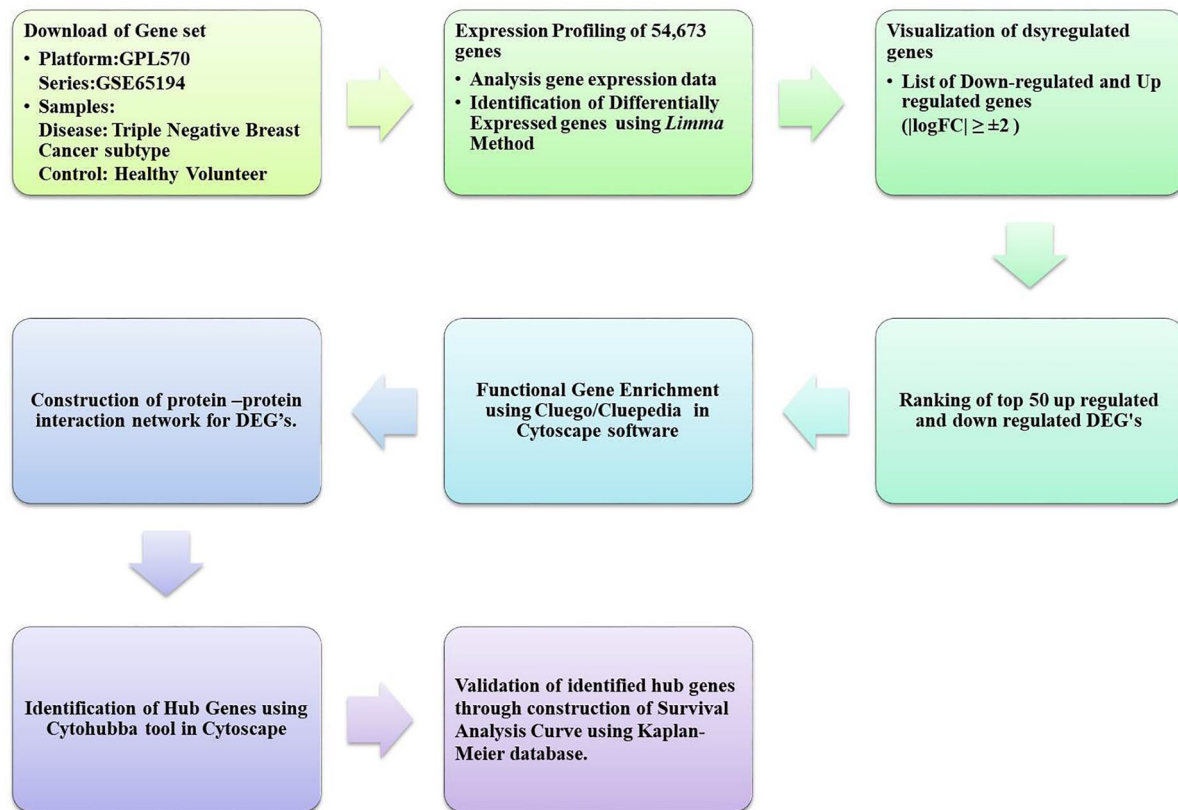


Figure 1. Flowchart to investigate potential signature biomarkers in triple-negative prognosis. The flowchart represents an overview of the steps followed for data collection and method implementation to conduct this study.

are expressed differentially in breast cancer patients. For instance, Chaun et al²² demonstrated CXCR4 and CXCL10 as predictive biomarkers for TNBC. Similarly, Zhai et al²³ reported the investigation of NUF2 and FAM83D as potential biomarkers in TNBC.

Thus, the study of biomarkers provides the potential for developing diagnostic tools and enhancing the efficiency of health care therapies.²⁴ To have adequate sensitivity and specificity for the diagnosis or prognosis of cancer, it is becoming more expected that signature biomarker analytes rather than single indicators will be needed.²⁵ Many studies have shown that signatures with an appropriate combination of many potential biomarkers can significantly increase predictive accuracy.^{26,27} Thus, the necessity for cancer-specific biomarker panels stems that can reveal the molecular complexity of human carcinogenesis even within single cancer sites.

This study has been carried out with an objective to identify differentially expressed genes (DEGs) using R package. Their bioinformatics-based analysis has been performed for identifying novel TNBC signature biomarkers. Assessment of gene ontology (GO) and pathway enrichment evaluation were performed using different tools. Protein-protein interaction (PPI) networks were established to integrate DEGs. Candidate key genes were identified using cytoHubba plugin in Cytoscape software. The clinical relevance of these genes in TNBC was

evaluated through survival analysis based on the Kaplan-Meier tool.

Methods

Information of microarray data set

The analysis of the microarray data set used in this study was performed on a data set retrieved from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database. The platform used for the TNBC data set samples was Affymetrix Human Genome U133A Array (GPL570 [HG-U133A]), and the file type was CEL files.

The “GSE65194” microarray data set was used in this study, which consisted of a total of 130 samples of breast cancer, including 41 samples of TNBC, 30 samples of HER2-enriched, 30 samples of Luminal B and 29 samples of Luminal A subtypes, as well as 11 samples of normal breast tissue and 14 TNBC cell lines.

For this study, a subset of 41 TNBC and 11 normal breast tissue samples were selected for further analysis. The selection of this subset was based on the principle of providing a diverse representation of breast cancer subtypes and normal samples, to conduct a comprehensive analysis of the gene expression data.

The pipeline of various tools used to conduct this study to predict the signature biomarkers in TNBC is illustrated in Figure 1.

Data set processing of differentially expressed genes

For bioinformatics analysis, R software (v4.1.2; <http://www.r-project.org>) was used. The DEGs in the gene expression profile of the data set was analyzed using GEO query, “dplyr” and “Limma” R package. The screening principle of DEGs was based on the log₂ Fold Change (log₂FC) and *P*-value set between targeted (TNBC patients) and control group (Normal Breast Tissue). The set filter conditions were adjusted to *P*-value < .01 and |log₂FC| ≥ ±2 as cut-off criteria to identify the DEG. The Benjamini-Hochberg (BH) method test was performed to investigate the “observed” sensitivity of the false discovery present in our test sample considering |log₂FC| ≥ 2.0 and adjusted *P* < .01 as cut-off criteria to be significant. The DEGs were visualized through volcano plot using ggplot2 R package.

Functional and pathway enrichment analysis

To investigate the functions of DEGs, functional enrichment analysis was performed using the ClueGo and CluePedia plugins for Cytoscape (3.9.1) tool (<https://cytoscape.org/download.html>). The ClueGO tool was used to decipher functionally grouped GO and pathway annotation network to understand their implication in categories of molecular function, Kyoto Encyclopedia of Genes and Genome (KEGG) and Reactome pathway. The relationship between each term was calculated using K statistic, and adjusted *P* < .05 was considered as a statistical significant difference.

Construction of protein-protein interaction network

Cytoscape software (version 3.9.1) (<https://cytoscape.org/download.html>) was used for the visualization of PPI networks generated by mapping DEGs by the Search Tool for the Retrieval of Interacting Genes (STRING) application. Parameters such as score >0.4 text mining, experiments, databases, and coexpression as well as species limited to “Homo sapiens” were used to create validatory PPI network.

Hub genes selection and analysis

To explore hubs genes in PPI networks, cytoHubba plugin of Cytoscape (3.9.1) tool (<https://cytoscape.org/download.html>) was applied on the networks that provided a user-friendly interface to identify important nodes present in our biological networks. cytoHubba possess 11 topological analysis methods for ranking nodes in a PPI network by their network features. For selection, we applied 2 topological methods, ie, Degree and Bottleneck parameter at cut-off of 2 for electing overlapping hub genes from up-regulated and down-regulated PPI networks.

External verification of hub genes expression level

To verify the alteration in the expression profiling of the key hub genes, Gene Expression Profiling Interactive Analysis (GEPIA2) platform was used. The tool validates the expression level by comparing transcriptomic data from The Cancer Genome Atlas (TCGA), TCGA normal, and the Genotype-Tissue Expression (GTEx) database. The TCGA data sets accessed through the GEPIA2 platform (<http://gepia.cancer-pku.cn>) can be independently verified and accessed at <http://portal.gdc.cancer.gov/>. The following parameters were set for the box plot expression: |log₂FC| cut-off=1, *P*-value cut-off=0.01, log scale=yes, jitter size=0.4, and match TCGA normal data. The box plot was used to illustrate the link between the gene expression and the samples.

The gene expression profiling of hub genes by box plot analysis was also performed on the data set “GSE65194” employed in this study, and the results were compared with those obtained from GEPIA2. This comparison allows for the identification of genes that are consistently differentially expressed across multiple data sets, which increases the confidence in the results. The comparison of results from GEPIA2 and GSE65194 data sets allows to ensure that the identified key hub genes are truly differentially and biologically relevant and can be used for further research and understanding of the disease.

Survival analysis

To study the clinical outcome of hub genes, predictive studies for overall survival (OS) were carried out using the Kaplan-Meier database (<http://kmplot.com/analysis/>), an online database that includes gene expression and clinical data. This database contains information on lung cancer, ovarian cancer, gastric cancer, and breast cancer. We used the Kaplan-Meier survival plot to assess the OS of breast cancer patients. The survival plots were obtained by uploading 8 genes into the database. Survival analysis for each gene was conducted using Cox proportional hazard regression, log-rank *P*-value, and the BH technique for multiple hypothesis testing.

Results

Identification of differentially expressed genes in triple-negative breast cancer

On analyzing the gene expression microarray data of GSE65194, up-regulated (2263) and down-regulated (980) genes were obtained from the data set of 54674 genes using *limma* software package. For visualizing overall gene expression of differential genes, volcano map was generated using log₂FC score and log₁₀ *P*-values (ggplot2 package) in R software, where the dots are in 2 colors. In this visualization, up-regulated and down-regulated genes are in distinct colors, namely, red and blue, which have been denoted, respectively (Figure 2).

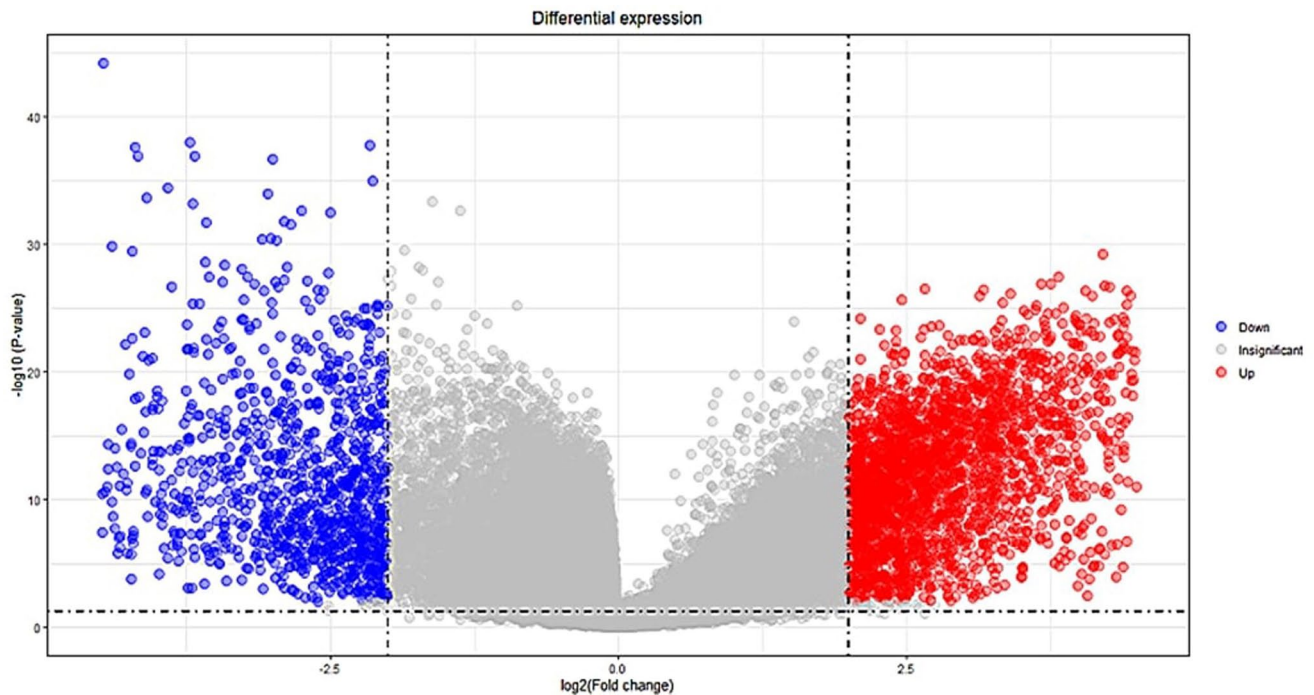


Figure 2. Volcano map of Differentially Expressed Genes: Volcano plot of differentially expressed genes between triple-negative breast cancer (TNBC) and Health Patient with $|\log_{2}FC| \geq 2$, P -value threshold = 0.01 and adjusted P -value (q -value) threshold $< .0001$, where the red dots represented on the right-side panel are the genes that are up-regulated in the breast cancer, the blue dots represented on the left-side panel are the genes that are down-regulated in breast cancer, and the gray dots represent the genes that are not significantly differentiated.

The BH method was used to adjust the raw P -values into a false discovery rate (FDR), $P < .01$ and $|\log_{2}FC| \geq 2$ were used as selection criteria. Thereafter, top 47 up-regulated DEGs and top 47 down-regulated DEGs were selected. The top DEGs are shown in Tables 1 and 2, and additional information of data has been provided in Supplementary Tables S1 and S2, representing the tabulated information of selected up-regulated and down-regulated DEGs.

ClueGO/CluePedia enrichment analysis

To investigate the potential biological function and molecular pathways that contribute to the pathogenesis of TNBC, DEGs were subjected to GO functional enrichment, *KEGG*, and *Reactome* pathway enrichment analysis. ClueGO/CluePedia plugin in Cytoscape software was used (Figure 3A and B). The enrichment analyses were carried out using hypergeometric test and kappa score ≥ 0.4 as primary criteria.

The up-regulated DEGs were found to be significantly enriched in molecular function for regulation of ubiquitin protein activities. The most significant terms of group were overviewed in the pie chart. It was seen that the 75% terms of up-regulated DEGs were associated with the regulation of ubiquitin protein ligase activity and 12.5% terms in histone kinase activity and regulation activity of serine/threonine kinases (Supplementary Figure S1 (a)). Correspondingly, *KEGG* and *Reactome* pathway analysis (Supplementary Figure

S1 (b) and (c)) revealed that 60% terms of up-regulated DEGs were significantly engaged in cell cycle pathway, PR-mediated oocyte maturation and oocyte meiosis, 20% gene in pyrimidine metabolism, and remaining genes were contributing to controlling cyclin A/B1/B2 associated events during G2/M transition, sumoylation of DNA replication proteins, and p53 signaling pathway. Genes such as *AURKA*, *BUB1*, and *BIRC5* were identified to be involved in different stages of the cell cycle; similarly, *KIF2C* gene was seen to possess a role in controlling microtubule dynamics, whereas genes such as *CDC20* and *UBE2* were observed to influence the ubiquitin activities. Supplementary Table S3 (A) to (C) provides tabulated information of all the up-regulated DEGs functional enrichment.

Down-regulated DEGs molecular function (Figure 4A) was found to be engaged in pathway regulation of protein dimerization activity, protein domain-specific binding, nucleotide and nucleoside phosphate binding, and identical protein binding (Figure 4B and C). In addition to these, the genes were enriched in the peroxisome proliferator-activated receptor (PPAR) signaling pathway, RHO GTPase cycle, pyruvate metabolism, and so on.

It was revealed that the genes like *ARHGEF15*, *OPHN1*, and *RAPGEF3* were involved in the GTPase activator activity, 25.75% terms of down-regulated DEGs genes contribute to the metabolism of vitamins and cofactors, whereas other genes like *PDE2A*, *ACACB*, and *VERGFD* were involved in MAPK signaling pathway, RHOJ GTPase cycle and PPAR signaling

Table 1. List of top 47 ranks up-regulated DEGs.

ID	LOGFC	P	FDR (AP)	GENES	GENE NAME
209773_s_at	8.03836	4.58E-41	1.24E-37	RRM2	Ribonucleotide reductase regulatory subunit M2
201291_s_at	7.66269	8.94E-41	2.42E-37	TOP2A	DNA topoisomerase alpha
218009_s_at	7.0028	2.86E-40	7.73E-37	PRC1	Protein regulator of cytokinesis 1
210052_s_at	6.87315	4.99E-40	1.35E-36	TPX2	TPX2 microtubule nucleation factor
201292_at	6.83254	2.06E-39	5.57E-36	TOP2A	DNA topoisomerase alpha
207828_s_at	6.65425	1.08E-37	2.93E-34	CENPF	Centromere protein F
219918_s_at	6.96352	1.62E-37	4.37E-34	ASPM	Assembly factor for spindle microtubule
202954_at	6.63982	5.10E-37	1.38E-33	UBE2C	Ubiquitin-conjugating enzyme E2C
213226_at	4.96515	9.00E-37	2.42E-33	CCNA2	Cyclin A2
218782_s_at	7.04088	1.01E-36	2.71E-33	ATAD2	ATPase family AAA domain containing 2
228273_at	5.8842	1.21E-36	3.27E-33	PRR11	Proline rich 111
201663_s_at	6.90789	2.65E-36	7.13E-33	SMC4	Structural maintenance of chromosome 2
209408_at	6.48523	4.41E-35	1.19E-31	KIF2C	Kinesin family member 2C
202870_s_at	7.21826	5.08E-35	1.37E-31	CDC20	Cell Division cycle 20
203358_s_at	5.59236	4.32E-34	1.16E-30	EZH2	Enhancer of zeste 2 polycomb repressive complex 2 subunit
201695_s_at	6.02508	5.13E-34	1.38E-30	PNP	Purine nucleoside phosphorylase
218039_at	6.48278	6.69E-34	1.80E-30	NUSAP1	Nucleolar and spindle associated protein 1
204444_at	5.32595	8.79E-34	2.36E-30	KIF11	Kinesin family member 11
219787_s_at	5.25861	9.87E-34	2.65E-30	ECT2	Epithelial cell transforming 2
202779_s_at	5.70574	1.58E-33	4.24E-30	UBE2S	Ubiquitin-conjugating enzyme E2S
201890_at	6.40048	2.04E-33	5.47E-30	RRM2	Ribonucleotide reductase regulatory subunit M2
214710_s_at	6.036	6.12E-33	1.64E-29	CCNB1	Cyclin B1
223229_at	6.70473	6.41E-33	1.72E-29	UBE2T	Ubiquitin-conjugating enzyme E2T
203362_s_at	5.9518	7.25E-33	1.94E-29	MAD2L1	Mitotic arrest deficient 2 like 1
202705_at	6.16826	1.78E-32	4.77E-29	CCNB2	Cyclin B2
218542_at	5.31212	2.89E-32	7.74E-29	CEP55	Centrosomal protein 55
210639_s_at	5.38666	3.90E-32	1.04E-28	ATG5	Autophagy related 5
210559_s_at	6.36547	6.03E-32	1.61E-28	CDK1	Cyclin-dependent kinase 1
209464_at	4.59495	1.06E-31	2.83E-28	AURKB	Aurora kinase B
209642_at	5.30713	1.34E-31	3.59E-28	BUB1	BUB1 mitotic checkpoint serine/threonine kinase
213007_at	4.78874	1.57E-31	4.19E-28	FANC1	FA complementation group 1
226319_s_at	6.50911	2.07E-31	5.54E-28	ALYREF	Aly/REF export factor
202095_s_at	6.47966	3.38E-31	9.03E-28	BIRC5	Baculoviral IAP repeat containing 5
222608_s_at	5.75055	3.53E-31	9.42E-28	ANLN	Anillin actin binding protein
204825_at	5.9786	3.99E-31	1.06E-27	MELK	Maternal embryonic leucine zipper kinase
204962_s_at	6.13986	1.89E-30	5.04E-27	CENPA	Centromere protein A

(Continued)

Table 1. (Continued)

ID	LOGFC	P	FDR (AP)	GENES	GENE NAME
202613_at	5.03419	2.15E-30	5.72E-27	CTPS1	CTP synthase 1
218585_s_at	5.46961	2.47E-30	6.59E-27	DTL	Denticleless E3 ubiquitin protein ligase homolog
203214_x_at	6.20147	5.57E-30	1.48E-26	CDK1	Cyclin-dependent kinase 1
218349_s_at	4.20374	6.10E-30	1.62E-26	ZWILCH	Zwilich kinetochore protein
222740_at	4.81961	8.51E-30	2.27E-26	ATAD2	ATPase family AAA domain containing 2
211450_s_at	5.43724	9.66E-30	2.57E-26	MSH6	mutS homolog 6
204822_at	6.46714	1.70E-29	4.52E-26	TTK	TTK protein kinase
211762_s_at	4.73884	1.75E-29	4.64E-26	KPNA2	Karyopherin subunit alpha 2
202589_at	5.44127	2.39E-29	6.34E-26	TYMS	Thymidylate synthetase
204170_s_at	5.72354	2.63E-29	6.99E-26	CKS2	CDC28 protein kinase regulatory subunit 2
208079_s_at	5.42979	3.58E-29	9.50E-26	AURKA	Aurora kinase A
204092_s_at	5.37507	4.65E-29	1.23E-25	AURKA	Aurora kinase A
229128_s_at	7.73067	4.65E-29	1.23E-24	ANP32E	Acidic nuclear phosphoprotein 32 family member E

Table 2. List of top 47 ranks down-regulated DEGs.

ID	LOGFC	P	B	FDR (AP1)	GENE SYMBOL	GENE NAME
224061_at	4.46987461	5.94E-45	91.3257	7.33E-42	INMT	Indolethylamine N-methyltransferase
211565_at	6.36304759	1.52E-44	90.4416	1.87E-41	SH3GL3	SH3 domain containing GRB2 like 3, endophilin A3
236359_at	4.98314893	6.97E-42	84.6232	8.59E-39	SCN4B	Sodium voltage-gated channel beta subunit 4
232538_at	5.65002858	6.71E-41	82.4563	8.26E-38	-	-
220232_at	4.53268892	8.69E-40	79.9963	1.07E-36	SCD5	Stearoyl-CoA desaturase 5
216333_x_at	7.53215863	1.39E-39	79.5429	1.71E-36	TNXB	Tenascin XB
206093_x_at	7.3976163	1.98E-39	79.2037	2.43E-36	TNXA	Tenascin XA (pseudogene)
228618_at	3.71352097	1.08E-38	77.5672	1.32E-35	PEAR1	Platelet endothelial aggregation receptor 1
227106_at	2.15515744	1.83E-38	77.056	2.24E-35	TMEM198B	Transmembrane protein 19B (pseudogene)
204754_at	4.18930264	2.41E-38	76.7908	2.95E-35	HLF	HLF transcription factor, PAR bZIP family member
213451_x_at	7.28038202	2.51E-38	76.7518	3.07E-35	TNXA	Tenascin XA (pseudogene)
229177_at	4.16450659	1.18E-37	75.253	1.44E-34	C16orf89	Chromosome 16 open reading frame 89
226303_at	3.67720012	1.29E-37	75.1668	1.57E-34	PGM5	Phosphoglucomutase 5
213715_s_at	2.99546232	1.95E-37	74.7641	2.38E-34	KANK3	KN motif and ankyrin repeat domain 3
243689_s_at	6.11686746	4.23E-37	74.0141	5.16E-34	FRG1BP	FISH region gene 1 family member B (pseudogene)
221928_at	5.42959116	7.82E-37	73.4167	9.53E-34	ACACB	Acetyl-CoA carboxylase beta
229839_at	7.05960613	9.27E-36	71.0099	1.13E-32	SCARA5	Scavenger receptor class A member 5

(Continued)

Table 2. (Continued)

ID	LOGFC	P	B	FDR (AP1)	GENE SYMBOL	GENE NAME
233103_at	2.12847923	9.68E-36	70.9684	1.18E-32	-	-
210051_at	4.89322894	1.54E-35	70.5163	1.87E-32	RAPGEF3	Rap guanine nucleotide exchange factor 3
206742_at	5.23828534	2.88E-35	69.906	3.49E-32	VEGFD	Vascular endothelial growth factor D
1561754_at	5.78969636	3.06E-35	69.8438	3.72E-32		
237737_at	3.9096105	3.92E-35	69.6037	4.75E-32	ANKRD20A12P	Ankyrin repeat domain20 family member A12 (pseudogene)
234675_x_at	3.03714406	1.12E-34	68.5806	1.35E-31	-	-
236010_at	4.09026114	2.09E-34	67.9705	2.53E-31	-	-
1562755_at	5.62456897	5.87E-34	66.9577	7.11E-31	-	-
224159_x_at	3.69040256	6.52E-34	66.8554	7.88E-31	TRIM4	Tripartite motif containing 4
224012_at	4.94675803	1.22E-33	66.2446	1.47E-30	ANKRD20A1	Ankyrin repeat domain 20 family member A1
216153_x_at	2.75093532	2.19E-33	65.6665	2.65E-30	RECK	Reversion-inducing cysteine rich protein with Kazal motif
205883_at	5.37719229	2.30E-33	65.6184	2.78E-30	ZBTB16	Zinc finger and BTB domain-containing 16
43427_at	7.03540385	2.54E-33	65.5243	3.06E-30	ACACB	Acetyl-CoA carboxylase beta
205507_at	2.49445768	3.23E-33	65.2869	3.89E-30	ARHGEF15	Rho guanine nucleotide exchange factor 15
205384_at	4.58300665	9.99E-33	64.1793	1.20E-29	FXYP1	FXYP domain-containing ion transport regulator 1
211673_s_at	2.89379964	1.48E-32	63.797	1.77E-29	MOCSC1	Molybdenum cofactor synthesis 1
204134_at	3.57062886	1.74E-32	63.6377	2.08E-29	PDE2A	Phosphodiesterase 2A
65718_at	2.84190389	2.70E-32	63.2028	3.24E-29	ADGRA2	Adhesion G protein-coupled receptor A2
228854_at	7.29140473	3.91E-32	62.84	4.69E-29	-	-
227419_x_at	5.36433943	4.46E-32	62.7112	5.34E-29	PLAC9	Placenta associated 9
49452_at	7.88857109	8.28E-32	62.1013	9.92E-29	ACACB	Acetyl-CoA carboxylase beta
208609_s_at	5.06841005	1.01E-31	61.9098	1.20E-28	TNXB	Tenascin XB
213661_at	4.57393694	2.61E-31	60.9706	3.12E-28	PAMR1	Peptidase domain-containing associated with muscle regeneration 1
220725_x_at	3.01846022	3.09E-31	60.8056	3.69E-28	DNAH3	Dynein axonemal heavy chain 3
206323_x_at	3.08661046	3.86E-31	60.5865	4.61E-28	OPHN1	Oligophrenin 1
238701_x_at	2.96305494	4.62E-31	60.4108	5.50E-28	COLCA1	Colorectal cancer associated 1
210794_s_at	4.3921212	1.49E-30	59.257	1.77E-27	MEG3	Maternally expressed 3
1556364_at	4.21441323	3.30E-30	58.4716	3.93E-27	ADAMTS9-AS2	ADAMTS9 antisense RNA 2
238062_at	4.57077825	5.32E-30	58.0009	6.33E-27	GPIHBP1	Glycosylphosphatidylinositol anchored high-density lipoprotein binding protein 1
204482_at	3.58153905	2.30E-29	56.5551	2.74E-26	CLDN5	Claudin 5
206262_at	3.41611997	4.33E-29	55.9306	5.14E-26	ADH1C	Alcohol dehydrogenase 1C
205730_s_at	4.60411862	6.00E-29	55.6086	7.12E-26	ABLIM3	Actin binding LIM protein family member 3

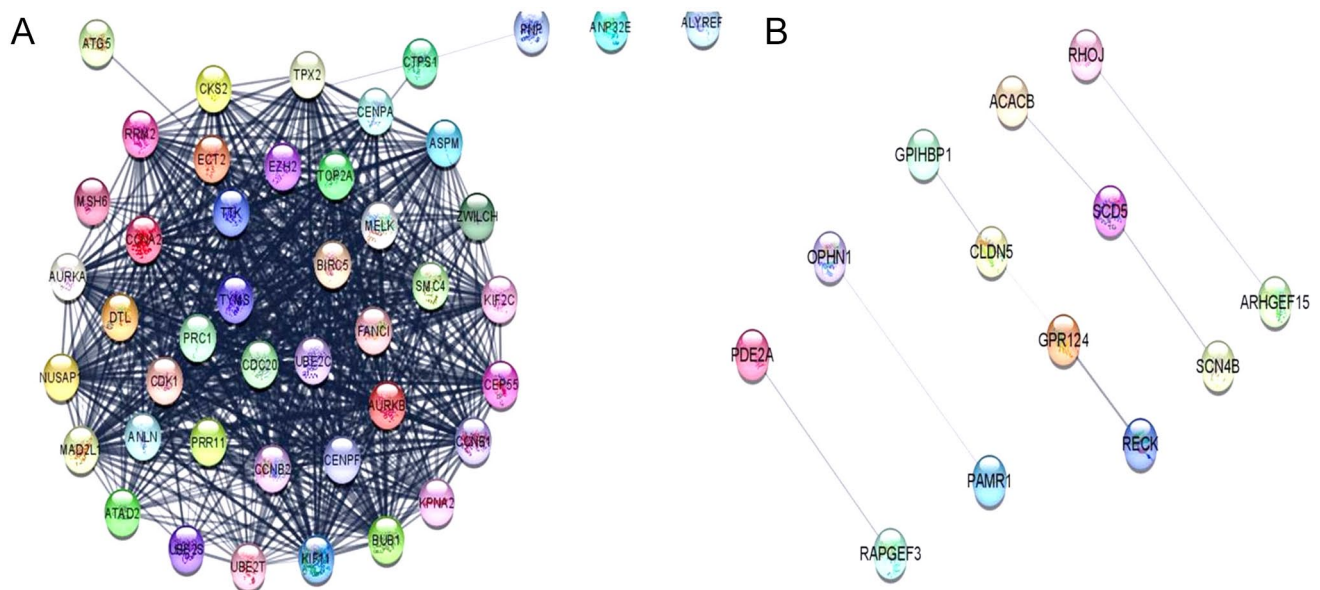


Figure 3. Schematic representation of protein-protein interaction (PPI) network analysis of differentially expressed genes established using Cytoscape software. Parameters set for analysis include degree cut-off=2, node score=0.2, k-score=2. Each node represents the protein, and edges represent the interaction of proteins. (A) The PPI network of up-regulated DEGs and (B) the PPI network of down-regulated DEGs.

pathway, and Rap1 signaling pathway, respectively. Supplementary Figure S2 (a) to (c) and Supplementary Table S4 (A) to (C) contain the tabulated data of all the down-regulated DEG's functional enrichment.

The results obtained from *ClueGO* enrichment clearly illustrates that majorly up-regulated DEG's contribute to regulating cell cycle and ubiquitin protein activities, whereas down-regulated DEG's were found to be involved in regulation of mitotic metaphase/anaphase transition, change in major signaling pathways, metabolic reprogramming, and so on.

Protein-protein interaction network construction and module analysis

To identify and understand the interaction among DEGs, PPI network was constructed using *STRING* database. Figure 5A and B represents the PPI network of up-regulated and down-regulated genes, respectively. The PPI network established a single network comprising of 37 nodes and 665 edges in case of up-regulated DEGs and 5 clusters comprising of overall 32 nodes and 8 edges for down-regulated genes at the combined score >0.4.

CytoHubba

To gain insights into the underlying biology of TNBC, which could be pivotal in governing various cellular processes and is often linked to critical functions and pathways, hub genes were identified. These hub genes, often described as a group of genes occupying central positions within biological networks or pathways due to their extensive connectivity and interactions

with other genes or proteins, were pinpointed using the degree and bottleneck algorithms with a cut-off value of 2.

From the pool of DEGs, 5 hub genes (TTK, CCNB1, BIRC5, Cytidine nucleotide triphosphate synthase 1 [CTPS1], thymidylate synthase [TYMS]) were identified among the up-regulated genes and 3 hub genes (SCD5, CLDN5, GPR124) among the down-regulated genes. Supplementary Tables S5 and S6 contain the cytoHubba results for up-regulated and down-regulated DEGs, respectively.

Among the down-regulated hub genes, there was a direct interaction observed between CLDN5 and GPR124. Similarly, among the up-regulated hub genes, direct connections were found among TTK, CCNB1, BIRC5, CTPS1, and TYMS, with CTPS1 having direct links to BIRC5 and TYMS. A comprehensive list of these hub genes is shown in Tables 3 and 4.

External verification of hub genes expression level

To verify the expression levels of identified hub genes in breast cancer samples, TCGA (<http://portal.gdc.cancer.gov/>) data set was analyzed using the GEPIA2 platform. The findings are represented by box plots illustrating the expression levels of the top 10 hub genes in breast invasive carcinoma (BRCA) samples compared with normal breast tissue samples, with 1085 tumor samples (num (T)=1085) and 112 normal samples (num (N)=112). For the overexpressed hub genes, the box plots show significantly higher median expression levels in tumor samples, suggesting their up-regulation in breast cancer and potential roles in cancer development. Conversely, the box plots for down-regulated hub genes display lower expression levels in

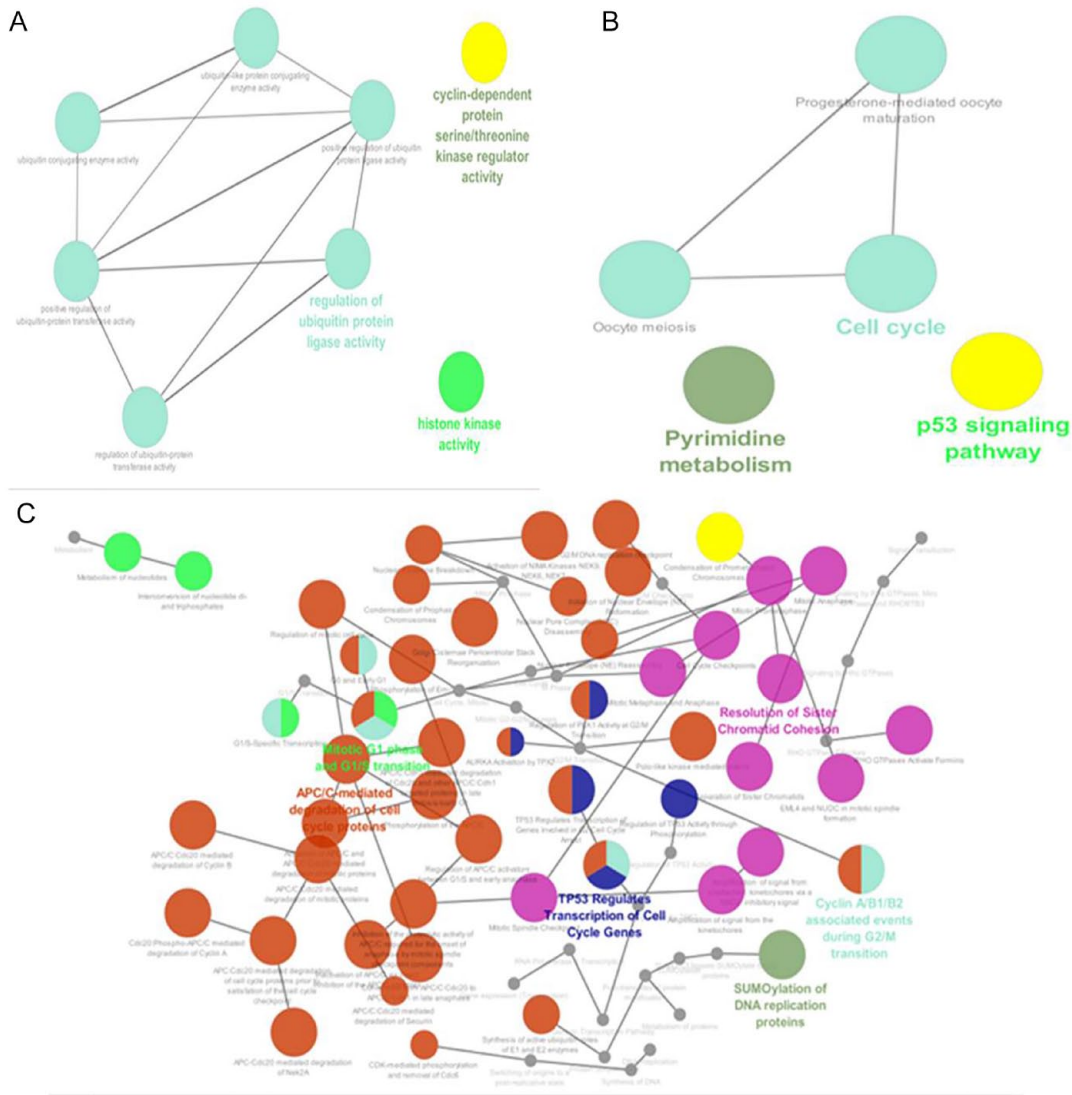


Figure 4. Illustration of functional enrichment analysis was visualized using ClueGO/CluePedia plugin in Cytoscape. The plugin provides an extensive enrichment analysis for (A) molecular pathway, (B) KEGG pathway, and (C) Reactome pathway of up-regulated genes. The network’s functional nodes and edges that are shared by DEGs with a kappa value of 0.4 describe the connectivity of the pathways. The enrichment shows only significant pathways ($P \leq .05$). Different color coding indicates different molecular pathways that were examined during the enrichment analysis of the discovered DEGs. The names of pathways are denoted by bold typefaces, which highlight the most important functional pathways.

tumor samples, indicating their under expression and possible function as tumor suppressors as illustrated in Figure 6.

A similar pattern of alteration in the expression of the hub gene was reflected through the box plot of gene expression profile comparison in the selected data set (GSE651094) used in this study (num (T)=55, num (H)=11), as illustrated in Figure 7.

Survival Analysis

To evaluate the prognostic roles of the potential hub genes for TNBC, the *Kaplan-Meier* plotter analysis platform was applied. A total of 1880 TNBC patient’s data were available

for the analysis of OS. The results demonstrated that increased expression of TTK (HR (Hazard Ratio) =2.18 [1.69-2.82], logrank $P=1.2e-09$), CCNB1 (HR=1.64 [1.36-1.99], logrank $P=2.6e-07$), BIRC5 (HR=1.7 [1.4-2.05], logrank $P=4.7e-08$), CTPS1 (HR=1.39 [1.15-1.68], logrank $P=.00063$), and TYMS (HR=1.38 [1.14-1.67], logrank $P=.00084$) are associated with a lower survival rate of the patients than its low expression. In contrast, TNBC patients showed a lower likelihood of survival as a result of declining SCD5 expression level (HR=0.96 [0.79-1.16], logrank $P=.65$). Likewise, a poor OS rate was observed to be associated with decreased CLDN5 expression (HR=0.89 [0.74-1.04], logrank $P=.24$), suggesting these genes as

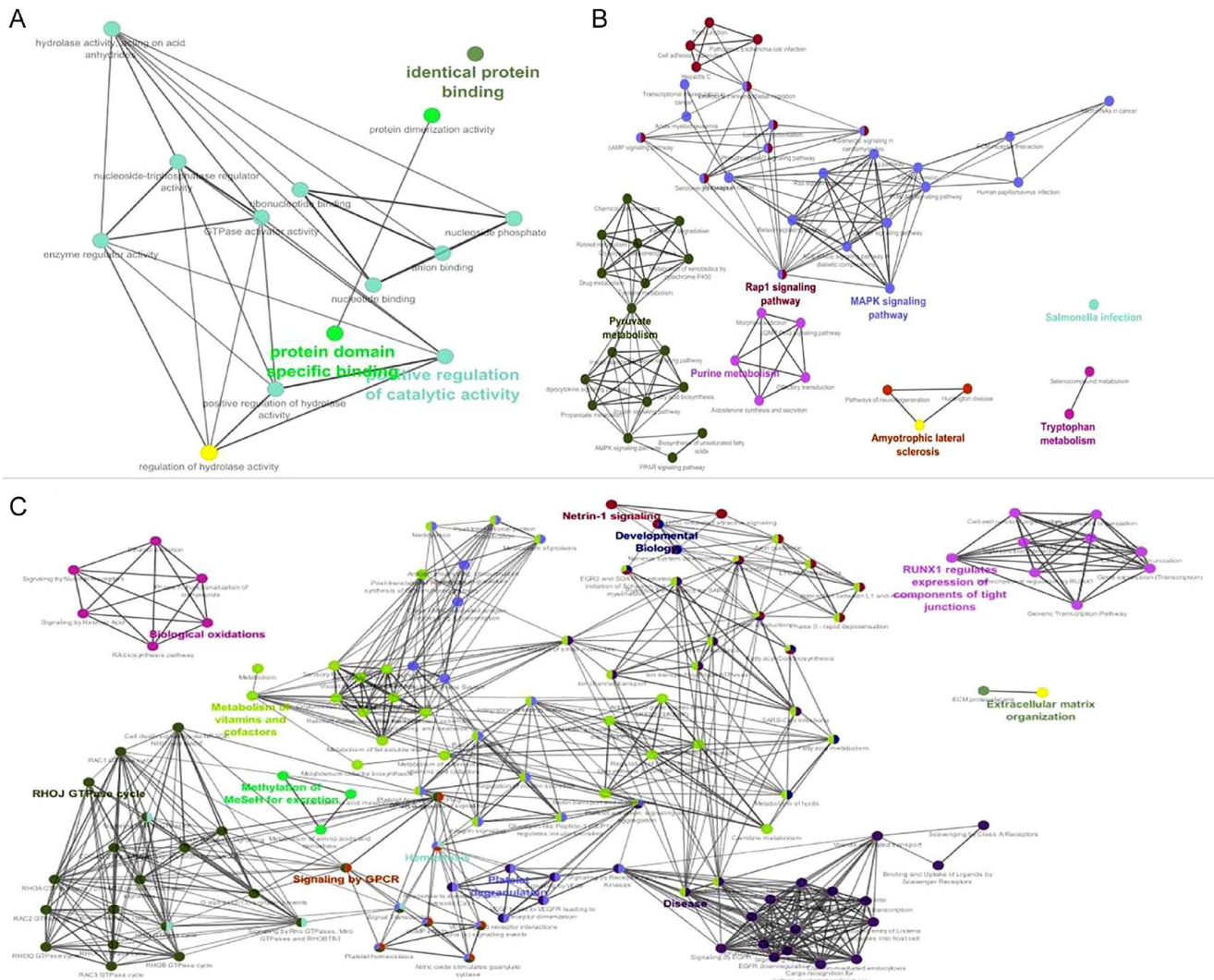


Figure 5. Illustration of Functional enrichment analysis was visualized using ClueGO/CluePedia plugin in Cytoscape. The plugin delivers a comprehensive enrichment analysis for down-regulated DEGs including (A) molecular pathway, (B) KEGG pathway, and (C) Reactome pathway of down-regulated genes. The network's functional nodes and edges that are shared by DEGs with a kappa value of 0.4 describe the connectivity of the pathways. The enrichment shows only significant pathways ($P \leq .05$). Different color coding indicates different molecular pathways that were examined during the enrichment analysis of the discovered DEGs. The names of pathways are denoted by bold typefaces, which highlight the most important functional pathways.

Table 3. Tabulated data of HUB genes identified from up-regulated DEGs.

UP-REGULATED GENES	NODE NAME	SYMBOL	DEGREE	BOTTLE NECK
	9606.ENSP00000301633	BIRC5	36.0	3.0
	9606.ENSP00000256442	CCNB1	37.0	2.0
	9606.ENSP00000315644	TYMS	36.0	3.0
	9606.ENSP00000358813	TTK	37.0	2.0
	9606.ENSP00000361704	CTPS1	3.0	2.0

Table 4. Tabulated Data of HUB genes identified from down-regulated DEGs.

DOWN-REGULATED GENES	NODE NAME	SYMBOL	DEGREE	BOTTLE NECK
	9606.ENSP00000406367	GPR124	2.0	2.0
	9606.ENSP00000385477	CLDN5	2.0	2.0
	9606.ENSP00000316329	SCD5	2.0	3.0

potential therapeutic targets for patients with TNBC. The survival analysis of GRP124 gene (HR=0.99 [0.82-1.2], logrank $P=0.93$) was not correlated with OS.

Discussion

In this study, the potential signature biomarkers in the breast cancer pathogenesis have been deciphered using bioinformatics strategy. A set of 2263 up-regulated and 980 down-regulated DEGs were identified between breast cancer and normal samples. The panel of up-regulated DEGs was found to be enriched in “Cell cycle,” “Oocyte meiosis,” “Regulation of ubiquitin protein ligase activities,” and so on. Cell cycle is an evolutionarily conserved and fundamental mechanism for growth and proliferation.²⁸ Its dysfunction is a defining feature of human cancer. Previous studies have indicated correlation of cell cycle with the outcome of breast cancer.²⁹ These dysregulations in cell cycle have been consequence of breast cancer metastasis.^{29,30} Several cell cycle-related genes including CCNB1 and CDK1—discovered as dysregulated genes in this study—have been implicated in the onset and progression of cancer.³¹ Likewise, abnormal regulation of meiosis and oocyte maturation has also been linked to affect cell cycle.^{32,33} A bioinformatics analysis reported by Wu et al shows that PR-mediated oocyte maturation and oocyte meiosis are associated with the pathogenesis of breast cancer.³⁴ A major regulator of protein ubiquitination, ubiquitin protein ligase (E3) functions to determine the fates of ubiquitin proteins. Aberrant functioning of E3 ubiquitin ligases results in altering the biological behavior of cancer cells, including breast cancer metastasis.³⁵ As a result, regulation of ubiquitin protein ligase activity demonstrates a critical role in metastasis of breast cancer. It is found that ubiquitin protein ligase exhibits its role in invasion, migration, angiogenesis, and immunity response.³⁶ In addition, these pathways have been acknowledged in the pathogenesis of prostate cancer and renal cell carcinoma.^{37,38} Similarly, down-regulated DEGs were found to be involved in “MAPK signaling pathways,” “Rap1 signaling pathway,” “PPAR signaling pathway,” “RHO GTPase cycle,” and so on. Among these pathways, MAPK pathway is primarily responsible for regulating the function of other proteins through phosphorylation.³⁹ Rap1 signaling regulates integrin or cadherin-mediated cell adhesion, expression levels of proteases, and cytoskeletal changes, which have been linked to cancer cell proliferation, development, and metastasis of cancer.⁴⁰ The PPAR is a nuclear hormone receptor observed to participate in cell proliferation

and differentiation, regulate immune response, and promote endothelial NO synthase (eNOS) activation in breast cancer.⁴¹ Other important pathway is the RHO GTPase cycle, which plays a critical role in breast cancer initiation, development, and metastasis. The Rho GTPase activation drives progression of cell through affecting metastatic cascade.⁴²

To detect highly connected protein nodes from PPI network, the *cytoHubba* plugin tool revealed SCD5, CLDN5, GPR124, TTK, CCNB1, BIRC5, CTPS1, TYMS genes as hub genes.^{43,44} The SCD5 is an endoplasmic reticulum-resident integral membrane protein that catalyzes the formation of monounsaturated fatty acids from saturated fatty acid. Down-regulation of SCD5 expression can be found to be related to more aggressive breast cancer phenotypes, such as high histological grade, late stage, and HER2 overexpression.⁴⁵ Claudin-5 is a protein member of the tight junction that is expressed in endothelial and epithelial cells.⁴⁶ It exhibits a critical role in maintaining homeostasis within the tissue environment. Overexpression of this protein is found to promote carcinogenesis and metastasis through increasing invasion and survival of cancer cells. In particular, previous studies revealed differential expression of claudins in cancers. High levels of Claudin-5 have been reported in ovarian, prostate, lung, and breast cancer; a “claudin-low” type has been found in TNBC.^{46,47} To date, only few studies have been addressed that corroborate the down-regulation of claudin-5 genes (similar to our finding) in breast cancer, especially TNBC.^{48,49} The TTK has been reported to be an attractive therapeutic target. It plays a key role in the spindle assembly checkpoint, mitotic regulation, and regulation of cell division processes, and it is up-regulated in a variety of cancers, such as bladder, esophagus, breast, lung, prostate, and anaplastic thyroid. Overexpression of TTK facilitates genomic instability in cancer cell proliferation and invasion. AI-Ejeh et al⁵⁰ found that elevation in the level of TTK protein in aggressive cancer leads to poor survival. Similarly, earlier TNBC studies also found that aberrant TTK expression in TNBC is an event that is significantly associated with an elevated risk of relapse and docetaxel resistance.^{51,52} These findings are further corroborated by this study's observation. In addition, it was discovered through survival analysis that overexpression of TTK gene imparts a negative effect on patients' chances of surviving. The CCNB1, a checkpoint protein, plays a significant role in apoptosis, chemo-resistance, and epithelial mesenchymal transitions to cancer cells.⁵³ Overexpression of cyclin B1 has been reported to

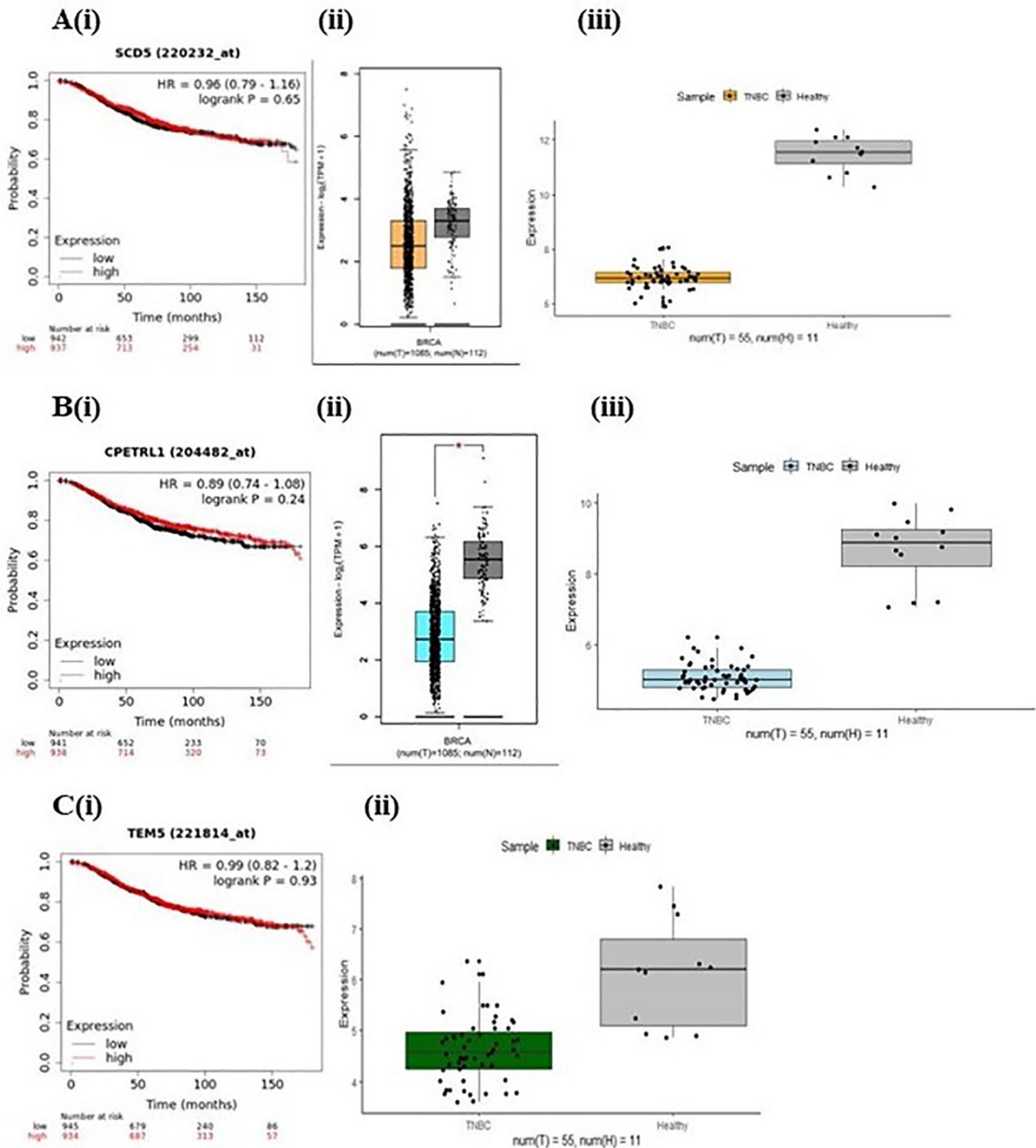


Figure 6. Verification of expression level of hub genes and survival analysis of up-regulated genes. Part (A) represents the survival graph of each hub gene. Line color in black represents the survival of the patient when the expression of the genes is low; red line below represents the probability of the survival of the patient when genes with high expression. Part (B) represents the box plot of 2 hub genes constructed using TCGA and GTEx expression data. Gene Expression Profiling Interactive Analysis (GEPIA) was performed. Different color code boxes represent the breast cancer tissue group, gray was the normal tissue group, and asterisk represented $P < .01$. The dots represented expression in each sample. Part (C) of figure represents the validity expression of 8 hub genes using test and control samples of GSE65194 data set. Different color code boxes represent the TNBC samples, and gray represents the normal sample. The dots represented expression in each sample.

be associated with many cancers, such as colorectal cancer, gastric cancer, pancreatic carcinoma, and lung carcinoma.⁵⁴ For instance, Aaltonen et al⁵⁵ demonstrated the correlation of

overexpression of the cyclin B1 with an aggressive phenotype and association with shorter OS and metastasis-free survival in breast cancer patients. A similar finding has been reported by

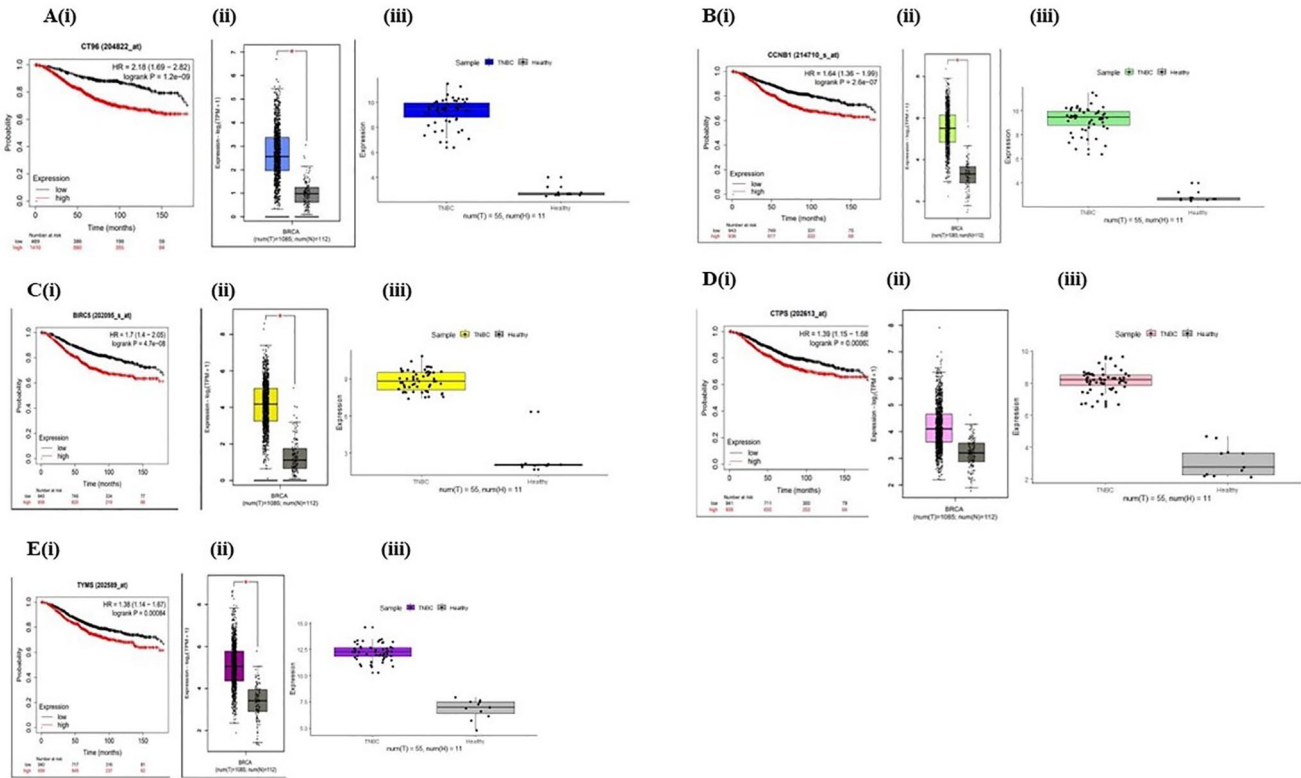


Figure 7. Verification of expression level of hub genes and survival analysis of down-regulated genes. In Part (A), each case represents the survival graph of each hub gene. Line color in black represents the survival of the patient when the expression of the genes is high; red line below represents the probability of the survival of the patient when genes with low expression. Part (B) represents the box plot of 2 hub genes constructed using TCGA and GTEx expression data. Gene Expression Profiling Interactive Analysis (GEPIA) was performed. Different color code boxes represent the breast cancer tissue group, gray was the normal tissue group, and asterisk represented $P < .01$. The dots represented expression in each sample. Part (C) of figure represents the validity expression of 8 hub genes using test and control samples of “GSE65194” data set. Different color code boxes represent the TNBC samples, and gray represents the normal sample. The dots represented expression in each sample.

Ding et al,⁵⁶ in that high level of CCNB1 was closely associated with hormone therapy resistance, poor recurrence-free survival, disease-free survival, and distant metastasis-free survival of ER+ breast cancer patients. Preliminary findings from this study show that CCNB1 is overexpressed in TNBC patients compared with healthy controls and may act as a potential biomarker for the disease, which is also laying the framework for future drug development. The CTPS1 is a CTP synthase that plays a critical role in DNA synthesis. Previous literature illustrates that higher CTPS1 gene expressions have association with worsen clinic-pathologic features such as larger cancer size, higher histological grade, and lympho-vascular invasion.⁵⁷ The results of our analysis also support the same trend, as TNBC patients with the higher expression of CTPS1 gene had poor OS rate. Few studies to date have evaluated its role in cancer development and progression.⁵⁸ The BIRC5 is a cell death preventing protein that is important for cell division in normal and cancerous cells. The BIRC5 is found during carcinogenesis in various cancer types. Treatment that targets BIRC5 has been increasingly noticed as a promising therapeutic strategy.⁵⁹ However, the detailed expression pattern, potential function, prognostic value, and drug

interaction network of BIRC5 remain largely unclear in breast cancer. The BIRC5 when overexpressed in breast cancer patients is responsible for a worse survival.⁶⁰ The TYMS is a key rate-limiting enzyme essential for the synthesis of thymidylate biosynthesis. This enzyme transfers tetrahydrofolate as methyl donor acid and catalyzes the conversion of methylation from deoxyuridine-monophosphate (dUMP) to deoxythymidine monophosphate (dTMP), which, on further phosphorylating, forms a triphosphate (dTTP), one of 4 the precursors used to synthesize DNA.⁶¹ The TYMS plays a crucial role in DNA synthesis and repair.⁶² Recent studies demonstrate that increase in expression level of TYMS has been observed in the development of several malignancies such as prostate cancer and lung cancer.⁶³ The GPR124 was discovered to be down-regulated in patients with TNBC through differential expression. Furthermore, the survival study showed no statistically significant link between the patients’ OS, demonstrating a poor prognostic indicator for the triple-negative breast subtype.

Our in silico investigations revealed BIRC5, CCNB1, and TTK as key signature genes that were observed to be significantly enriched in numerous biological pathways. Compared with normal breast tissue, these 3 genes in breast cancer samples

were found to be overexpressed, and the alternation of these genes was substantially connected with a poor survival outcome in breast cancer patients. Notably, these genes, in particular, have consistently ranked among the top dysregulated genes in other cancers like lung adenocarcinoma,¹⁹ pancreatic ductal adenocarcinoma,⁶⁴ and esophageal squamous cell carcinoma.⁶⁵ Thus, the discovery of these parallels in TNBC and other cancers suggests a fundamental role for these genes in the development of cancer and strengthens their potential as therapeutic targets.⁶⁶

Taken together, in this study, the primary objective was to thoroughly examine the most probable candidate genes associated with breast cancer. The identification of these uncovered hub genes holds great promise in elucidating the intricate molecular mechanisms driving the development of TNBC. Beyond shedding light on the disease's underlying processes, these discoveries may open up exciting new avenues for therapeutic interventions.

Fortifying the findings involved drawing from a wealth of examples in the existing scientific literature. These examples serve to underscore the substantial role that candidate genes play in the genesis of cancer. Consequently, the study provides a compelling body of evidence that not only enhances the understanding of breast cancer but also paves the way for promising directions in future breast cancer research.

Author Contributions

S. H. contributed to formal analysis, investigation, writing. S. P. contributed to conceptualization, validation, supervision, Editing and review. M. C. contributed to conceptualization, validation, supervision. V. P. contributed to conceptualization, validation, supervision, writing, editing.

ORCID iD

Veena Puri  <https://orcid.org/0000-0002-0639-5897>

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

REFERENCES

- Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71:209-249. doi:10.3322/caac.21660
- Global health estimates: leading causes of death. Accessed April 7, 2022. <https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghle-leading-causes-of-death>
- Yin L, Duan JJ, Bian XW, Yu SC. Triple-negative breast cancer molecular subtyping and treatment progress. *Breast Cancer Res.* 2020;22:61. doi:10.1186/s13058-020-01296-5
- Wahba HA, El-Hadaad HA. Current approaches in treatment of triple-negative breast cancer. *Cancer Biol Med.* 2015;12:106-116. doi:10.7497/j.issn.2095-3941.2015.0030
- Almansour NM. Triple-negative breast cancer: a brief review about epidemiology, risk factors, signaling pathways, treatment and role of artificial intelligence. *Front Mol Biosci.* 2022;9:836417.
- Jiménez-Santos MJ, García-Martín S, Fustero-Torre C, Di Domenico T, Gómez-López G, Al-Shahrour F. Bioinformatics roadmap for therapy selection in cancer genomics. *Mol Oncol.* 2022;16:3881-3908. doi:10.1002/1878-0261.13286
- Kulasingam V, Diamandis EP. Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. *Nat Clin Pract Oncol.* 2008;5:588-599. doi:10.1038/ncponc1187
- Rode MP, Silva AH, Cisolotto J, Rosolen D, Creczynski-Pasa TB. MiR-425-5p as an exosomal biomarker for metastatic prostate cancer. *Cell Signal.* 2021;87:110113. doi:10.1016/j.cellsig.2021.110113
- Zhao H, Zhang Q, Chen H, Rahman MR, Faruquee HM. Integrated multi-omics approach identified molecular mechanism and pathogenetic processes of COVID-19 that affect patient with Parkinson's disorder. *Saudi J Biol Sci.* 2021;28:6939-6945. doi:10.1016/j.sjbs.2021.07.074
- Wu D, Rice CM, Wang X. Cancer bioinformatics: a new approach to systems clinical medicine. *BMC Bioinformatics.* 2012;13:71. doi:10.1186/1471-2105-13-71
- Brenner C. Applications of bioinformatics in Cancer. *Cancers (Basel).* 2019;11:1630. doi:10.3390/cancers11111630
- Gobin E, Bagwell K, Wagner J, et al. A pan-cancer perspective of Matrix Metalloproteases (MMP) gene expression profile and their diagnostic/prognostic potential. *BMC Cancer.* 2019;19:581. doi:10.1186/s12885-019-5768-0
- Liu S, Wang W, Zhao Y, Liang K, Huang Y. Identification of potential key genes for pathogenesis and prognosis in prostate cancer by integrated analysis of gene expression profiles and the cancer genome atlas. *Front Oncol.* 2020;10:809.
- Shih AJ, Menzin A, Whyte J, et al. Identification of grade and origin specific cell populations in serous epithelial ovarian cancer by single cell RNA-Seq. *PLOS ONE.* 2018;13:e0206785. doi:10.1371/journal.pone.0206785
- Sinha N, Kumar V, Puri V, et al. Urinary exosomes: potential biomarkers for diabetic nephropathy. *Nephrology (Carlton).* 2020;25:881-887. doi:10.1111/nep.13720
- Gupta A, Puri S, Puri V. Bioinformatics unmasks the maneuverers of pain pathways in acute kidney injury. *Sci Rep.* 2019;9:11872. doi:10.1038/s41598-019-48209-x
- Puri V, Arora A, Puri S. Identification of the potential molecular biomarkers of COVID-19 and the associated drugs: a bioinformatics approach. *J Adv Sci Res.* 2021:204-224.
- Shi T, Gao G. Identify potential prognostic indicators and tumor-infiltrating immune cells in pancreatic adenocarcinoma. *Biosci Rep.* 2022;42:BSR20212523. doi:10.1042/BSR20212523
- Fei H, Chen S, Xu C. Interactive verification analysis of multiple sequencing data for identifying potential biomarker of lung adenocarcinoma. *Biomed Res Int.* 2020;2020:8931419. doi:10.1155/2020/8931419
- Shi Q, Wang W, Jia Z, Chen P, Ma K, Zhou C. ISL1, a novel regulator of CCNB1, CCNB2 and c-MYC genes, promotes gastric cancer cell proliferation and tumor growth. *Oncotarget.* 2016;7:36489-36500. doi:10.18632/oncotarget.9269
- Feng Y, Spezia M, Huang S, et al. Breast cancer development and progression: risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes Dis.* 2018;5:77-106. doi:10.1016/j.gendis.2018.05.001
- Chuan T, Li T, Yi C. Identification of CXCR4 and CXCL10 as potential predictive biomarkers in Triple Negative Breast Cancer (TNBC). *Med Sci Monit.* 2020;26:e918281. doi:10.12659/MSM.918281
- Zhai X, Yang Z, Liu X, Dong Z, Zhou D. Identification of NUF2 and FAM83D as potential biomarkers in triple-negative breast cancer. *peerJ.* 2020;8:e9975. doi:10.7717/peerj.9975
- Califf RM. Biomarker definitions and their applications. *Exp Biol Med.* 2018;243:213-221. doi:10.1177/1535370217750088
- Sarhadi VK, Armengol G. Molecular biomarkers in cancer. *Biomolecules.* 2022;12:1021. doi:10.3390/biom12081021
- Muinao T, Deka Boruah HP, Pal M. Multi-biomarker panel signature as the key to diagnosis of ovarian cancer. *Heliyon.* 2019;5:e02826. doi:10.1016/j.heliyon.2019.e02826
- Tang H, Wang S, Xiao G, et al. Comprehensive evaluation of published gene expression prognostic signatures for biomarker-based lung cancer clinical studies. *Ann Oncol.* 2017;28:733-740. doi:10.1093/annonc/mdw683
- Matthews HK, Bertoli C, de Bruin RAM. Cell cycle control in cancer. *Nat Rev Mol Cell Biol.* 2022;23:74-88. doi:10.1038/s41580-021-00404-3
- Thu KL, Soria-Bretones I, Mak TW, Cescon DW. Targeting the cell cycle in breast cancer: towards the next phase. *Cell Cycle.* 2018;17:1871-1885. doi:10.1080/15384101.2018.1502567
- Königsberg R, Rögelsperger O, Jäger W, et al. Cell cycle dysregulation influences survival in high risk breast cancer patients. *Cancer Invest.* 2008;26:734-740. doi:10.1080/07357900801944864
- Shi Y, Sang X, Deng J, et al. Six cell cycle-related genes serve as potential prognostic biomarkers and correlated with immune infiltrates in hepatocellular carcinoma. *J Cancer.* 2023;14:9-23.
- He M, Zhang T, Yang Y, Wang C. Mechanisms of oocyte maturation and related epigenetic regulation. *Front Cell Dev Biol.* 2021;9:654028.

33. Bower JJ, Vance LD, Psioda M, et al. Patterns of cell cycle checkpoint deregulation associated with intrinsic molecular subtypes of human breast cancer cells. *NPJ Breast Cancer*. 2017;3:9-12. doi:10.1038/s41523-017-0009-7
34. Wu D, Han B, Guo L, Fan Z. Molecular mechanisms associated with breast cancer based on integrated gene expression profiling by bioinformatics analysis. *J Obstet Gynaecol*. 2016;36:615-621. doi:10.3109/01443615.2015.1127902
35. Callis J. The ubiquitination machinery of the ubiquitin system. *Arabidopsis Book*. 2014;12:e0174. doi:10.1199/tab.0174
36. Wang Y, Dai J, Zeng Y, Guo J, Lan J. E3 ubiquitin ligases in breast cancer metastasis: a systematic review of pathogenic functions and clinical implications. *Front Oncol*. 2021;11:752604.
37. Chen Z, Lu W. Roles of ubiquitination and SUMOylation on prostate cancer: mechanisms and clinical implications. *Int J Mol Sci*. 2015;16:4560-4580. doi:10.3390/ijms16034560
38. Corn PG. Role of the ubiquitin proteasome system in renal cell carcinoma. *BMC Biochem*. 2007;8:S4. doi:10.1186/1471-2091-8-S1-S4
39. Normanno N, Campiglio M, Maiello MR, et al. Breast cancer cells with acquired resistance to the EGFR tyrosine kinase inhibitor gefitinib show persistent activation of MAPK signaling. *Breast Cancer Res Treat*. 2008;112:25-33. doi:10.1007/s10549-007-9830-2
40. Looi C-K, Hii L-W, Ngai SC, Leong C-O, Mai C-W. The role of Ras-associated protein 1 (Rap1) in cancer: bad actor or good player? *Biomedicines*. 2020;8:334. doi:10.3390/biomedicines8090334
41. Augimeri G, Giordano C, Gelsomino L, et al. The role of PPAR γ ligands in breast cancer: from basic research to clinical studies. *Cancers*. 2020;12:2623. doi:10.3390/cancers12092623
42. Humphries B, Wang Z, Yang C. Rho GTPases: big players in breast cancer initiation, metastasis and therapeutic responses. *Cells*. 2020;9:2167. doi:10.3390/cells9102167
43. Fu W, Cheng G, Ding Y, Deng Y, Guo P. Identification of hub genes and its correlation with the prognosis of acute myeloid leukemia based on high-throughput data analysis. *Precis Radiat Oncol*. 2020;4:49-56. doi:10.1002/pro6.1089
44. Charitou T, Bryan K, Lynn DJ. Using biological networks to integrate, visualize and analyze genomics data. *Genet Sel Evol*. 2016;48:27. doi:10.1186/s12711-016-0205-1
45. Zhao W, Sun L, Li X, et al. SCD5 expression correlates with prognosis and response to neoadjuvant chemotherapy in breast cancer. *Sci Rep*. 2021;11:8976. doi:10.21203/rs.3.rs-117096/v1
46. Escudero-Esparza A, Jiang WG, Martin TA. Claudin-5 is involved in breast cancer cell motility through the N-WASP and ROCK signalling pathways. *J Exp Clin Cancer Res*. 2012;31:43. doi:10.1186/1756-9966-31-43
47. Tabariès S, Siegel PM. The role of claudins in cancer metastasis. *Oncogene*. 2017;36:1176-1190. doi:10.1038/ncr.2016.289
48. Popova OP, Kuznetsova AV, Bogomazova SY, Ivanov AA. Claudins as biomarkers of differential diagnosis and prognosis of tumors. *J Cancer Res Clin Oncol*. 2021;147:2803-2817. doi:10.1007/s00432-021-03725-0
49. Yang Y, Liu L, Fang M, Bai H, Xu Y. The chromatin remodeling protein BRM regulates the transcription of tight junction proteins: implication in breast cancer metastasis. *Biochim Biophys Acta Gene Regul Mech*. 2019;1862:547-556. doi:10.1016/j.bbagr.2019.03.002
50. Al-Ejeh F, Simpson PT, Sanus JM, et al. Meta-analysis of the global gene expression profile of triple-negative breast cancer identifies genes for the prognostication and treatment of aggressive breast cancer. *Oncogenesis*. 2014;3:e100. doi:10.1038/oncsis.2014.14
51. Maia AR, de Man J, Boon U, et al. Inhibition of the spindle assembly checkpoint kinase TTK enhances the efficacy of docetaxel in a triple-negative breast cancer model. *Ann Oncol*. 2015;26:2180-2192. doi:10.1093/annonc/mdv293
52. Györfy B, Bottai G, Lehmann-Che J, et al. TP53 mutation-correlated genes predict the risk of tumor relapse and identify MPS1 as a potential therapeutic kinase in TP53-mutated breast cancers. *Mol Oncol*. 2014;8:508-519. doi:10.1016/j.molonc.2013.12.018
53. Gavet O, Pines J. Progressive activation of CyclinB1-Cdk1 coordinates entry to mitosis. *Dev Cell*. 2010;18:533-543. doi:10.1016/j.devcel.2010.02.013
54. Ye C, Wang J, Wu P, Li X, Chai Y. Prognostic role of cyclin B1 in solid tumors: a meta-analysis. *Oncotarget*. 2016;8:2224-2232. doi:10.18632/oncotarget.13653
55. Aaltonen K, Amini RM, Heikkilä P, et al. High cyclin B1 expression is associated with poor survival in breast cancer. *Br J Cancer*. 2009;100:1055-1060. doi:10.1038/sj.bjc.6604874
56. Ding K, Li W, Zou Z, Zou X, Wang C. CCNB1 is a prognostic biomarker for ER+ breast cancer. *Med Hypotheses*. 2014;83:359-364. doi:10.1016/j.mehy.2014.06.013
57. van Kuilenburg AB, Meinsma R, Vreken P, Waterham HR, van Gennip AH. Identification of a cDNA encoding an isoform of human CTP synthetase. *Biochim Biophys Acta*. 2000;1492:548-552. doi:10.1016/s0167-4781(00)00141-x
58. Lin Y, Lin L, Fu F, et al. Quantitative proteomics reveals stage-specific protein regulation of triple negative breast cancer. *Breast Cancer Res Treat*. 2021;185:39-52. doi:10.1007/s10549-020-05916-8
59. Hamy AS, Bieche I, Lehmann-Che J, et al. BIRC5 (Survivin): a pejorative prognostic marker in stage II/III breast cancer with no response to neoadjuvant chemotherapy. *Breast Cancer Res Treat*. 2016;159:499-511. doi:10.1007/s10549-016-3961-2
60. Oparina N, Erlandsson MC, Fäldt Beding A, et al. Prognostic significance of BIRC5/survivin in breast cancer: results from three independent cohorts. *Cancers*. 2021;13:2209. doi:10.3390/cancers13092209
61. Wang X, Sun X, Du X, et al. Thymidylate synthase gene polymorphisms as important contributors affecting hepatocellular carcinoma prognosis. *Clin Res Hepatol Gastroenterol*. 2017;41:319-326. doi:10.1016/j.clinre.2016.10.012
62. Rahman L, Voeller D, Rahman M, et al. Thymidylate synthase as an oncogene: a novel role for an essential DNA synthesis enzyme. *Cancer Cell*. 2004;5:341-351. doi:10.1016/s1535-6108(04)00080-7
63. Song S, Tian B, Zhang M, et al. Diagnostic and prognostic value of thymidylate synthase expression in breast cancer. *Clin Exp Pharmacol Physiol*. 2021;48:279-287. doi:10.1111/1440-1681.13415
64. Zeng Y, Fan R. Identification and verification of CCNB1 as a potential prognostic biomarker by comprehensive analysis. *Sci Rep*. 2022;12:16153. doi:10.1038/s41598-022-20615-8
65. Hao M, Zhao T, Chen D, et al. Integrated bioinformatics analysis of potential biomarkers and candidate drugs of esophageal squamous cell carcinoma. *Med Data Min*. 2022;5:15. doi:10.53388/20220520015
66. Mamoor S. Differential expression of G-Protein coupled receptor 124 in cancers of the breast. *OSF Preprints*, May 11, 2021. doi:10.31219/osf.io/zy6hb