Overexpression of CCNE1 confers a poorer prognostic in triple negative breast cancer identified by bioinformatic analysis

Qianqian Yuan
Wuhan University Zhongnan Hospital

Lewei Zheng
Wuhan University Zhongnan Hospital

Yiqin Liao
Wuhan University Zhongnan Hospital

Gaosong Wu (wugaosongtj@163.com)
Zhongnan Hospital of Wuhan University https://orcid.org/0000-0001-9296-7080

Research

Keywords: Triple negative breast cancer, GEO, CCNE1, Prognosis

Posted Date: January 25th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-50005/v2

Version of Record: A version of this preprint was published on March 23rd, 2021. See the published version at https://doi.org/10.1186/s12957-021-02200-x.

Abstract

Background. Triple-negative breast cancer (TNBC) is a major subtype of breast cancer. Due to the lack of effective therapeutic targets, the prognosis is poor. In order to find an effective target, despite many efforts, the molecular mechanisms of TNBC are still not well understood which remain to be a profound clinical challenge.

Methods. To identify the candidate genes in the carcinogenesis and progression of TNBC, microarray datasets GSE36693 and GSE65216 were downloaded from Gene Expression Omnibus (GEO) database. The differentially expressed genes (DEGs) were identified, and functional and pathway enrichment analyses were performed using the Gene Ontology(GO) and Kyoto Encyclopedia of Genes and Genomes(KEGG) databases via DAVID. We constructed the protein-protein interaction network (PPI) and the performed the module analysis using STRING and Cytoscape. Then we reanalyzed the selected DEGs genes and the survival analysis was performed using cBioportal.

Results. A total of 140 DEGs were identified, consisting of 69 upregulated genes and 71 downregulated genes. Three hub genes were up-regulated among the selected genes from PPI and biological process analysis uncovered the fact that these genes were mainly enriched in p53 pathway and the pathways in cancer. Survival analysis showed that only CCNE1 may be involved in the carcinogenesis, invasion or recurrence of TNBC.

Conclusion. CCNE1 could confer a poorer prognostic in TNBC identified by bioinformatic analysis and play key roles in the progression of TNBC which may contribute potential targets for the diagnosis, treatment and prognosis assessment of TNBC.

Introduction

Breast cancer (BC) is considered the most commonly diagnosed malignant tumor and the leading cause of cancer-related death among worldwide females [1]. Triple negative breast cancer (TNBC) is one of the main tumor subtypes of BC, which lacks the expression of hormone receptors (estrogen/progesterone, ER/PR) and human epidermal growth factor receptor 2 (HER2) amplification, leading to a lack of effective treatment against the corresponding targets or receptors in addition to chemotherapy and radiotherapy. It accounts for 10–15% of all BC [2], and it usually appears in the form of high-grade invasive ductal carcinoma and has a higher early recurrences rate of between the first and third year of treatment, with the majority of deaths occurring within the first 5 years [3]. And it often manifests as distant metastases and is associated with poorer prognosis with comparison to other breast cancer subtypes [4, 5].

Due to the limited understanding of the pathogenesis, development, reproduction and molecular mechanisms of TNBC, the mortality rate remains high. Therefore, it is of great importance to illuminate the exact molecular mechanisms of carcinogenesis, proliferation and recurrence of TNBC in order to further develop more efficient diagnostic means as well as therapeutic options. In recent years, there have been some bioinformatics studies on TNBC which can help identify the differentially expressed genes (DEG) associated with TNBC carcinogenesis and progression and the functional pathways in which these genes participate, proving that integrated bioinformatical approaches is helpful for us to better exploring the potential mechanisms. Changes in expression levels can often reflect pathological conditions, and proteins encoded by these DEGs may be involved in different biological process, cellular behaviors, and molecular pathways during tumor progressions. In view of the diversity and complexity of TNBCs, omics technologies are necessary instruments to expand our understanding of TNBC subtypes [6]. As biomarkers, PR, ER and ERBB2/HER2 can be used as prognostic targets of breast carcinoma and are helpful to

suggest the most appropriate chemotherapeutic treatments. BRCA1 and BRCA2 gene mutations and homologous recombination (HR) pathways are involved in the DNA damage repair in TNBC [7].

Transcriptomic analysis showed that different genes such as MKI67, TOP2A, EGFR were overexpressed in TNBC, providing different treatment direction for TNBC subtypes [6]. In this study, two microarray datasets (GSE36693 and GSE65216) containing both TNBC and non-TNBC samples were retrieved from Gene Expression Omnibus (GEO) and downloaded for further acquiring DEGs .The DEGs in the two datasets above were obtained using the GEO2R online tool in their respective data sets and then using the Venn chart software to obtain the intersection. The Database for Annotation, Visualization and Integrated Discovery (DAVID) was used to analyze these DEGs including Gene Ontology (GO) and Kyoto Encyclopedia of Gene and Genome (KEGG) pathways, the former part containing biological process (BP), cellular component (CC), molecular function (MF). We constructed protein-protein interaction (PPI) network using STRING (Search Tool for the Retrieval of Interacting Genes), a convenient online software, and then applied Cytoscape as well as Molecular Complex Detection (MCODE) for extra analysis of DEGs in order to identify some hub-genes. Subsequently, re-analysis of important genes by GO and KEGG pathway enrichment were performed. Moreover, Gene Expression Profiling Interactive Analysis (GEPIA) was utilized to furtherly validated the DEGs expression between TNBC tissues and non-TNBC tissues. Then the survival of hub-genes was analyzed by cBioPortal.

Materials And Methods

Dataset selection

NCBI-GEO is considered as a public dataset containing countless microarray information [9], from which the gene expression profile of GSE36693 and GSE65216 in TNBC and non-TNBC tissues were downloaded. And the GSE36693 dataset contained 21 TNBC and 66 non-TNBC samples. GSE65216 contained 55 TNBC and 109 non-TNBC samples.

Data processing and identification of DEGs

DEGs between TNBC samples and non-TNBC specimen were identified via GEO2R website (http://www.ncbi.nlm.nih.gov/geo/geo2r) with cut-off levels of log2 (fold-change)>2 and adjust P value <0.01 [10]. Then, through Venn software, the intersection of the unsettled data in TXT format was taken and graphed online to identify common DEGs in the two databases.

GO and KEGG enrichment analyses of DEGs

GO analysis is a widely-used approach to make a definition of genes and their downstream product to identify unique biological properties of high-throughput information [11]. KEGG is a collection of databases that manage genomes, diseases, biological pathways and relative chemical materials [12]. And the online tool, DAVID (http://david.ncifcrf.gov), is utilized to integrate biological data and provides a complete series of functional annotation information of genes and their proteins for researchers to obtain biological data [13]. In this study, DAVID was used to perform enrichment analysis of DEG's BP, MF, CC and its pathways (P<0.05).

PPI network and module analysis

The online tool STRING (http://string-db.org) can be used to construct a PPI network [14]. Then, the results of STRING were further analyzed and visualized by Cytoscape to explore the possible association of these DEGs, with

maximum number of interactors = 0 and confidence level≥0.4 [15]. The MCODE of Cytoscape is an APP based on topological structure for clustering some known networks to find areas of their dense connectivity for further bioinformatic uses [16]. Thus, MCODE was utilized to check modules of the PPI network (MCODE scores>5, degree cutoff=2, max. Depth=100, k-core=2, and node score cutoff=0.2).

RNA sequencing expression of core genes

GEPIA is a network tool for cancer and normal gene expression profiling and interactive analyses to conduct differential expression analysis and survival analysis of genes of interest [17]. GEPIA website was applied to perform RNA expression analysis on these DEGs through TCGA data samples to verify.

Survival analysis

The cBioPortal online platform provides a visual analysis instrument for interactive exploration of diverse cancer genome datasets [18, 19], allowing Users to perform survival analysis based on DNA mutation data and CNA data, and visually displaying the patient's overall survival (OS) and distant free survival (DFS) results in the form of Kaplan-Meier diagrams [20]. Kaplan-Meier curve was drawn through the cBioPortal online platform to analyze the OS of hub genes. We then calculate the hazard ratio and log rank P value with 95% confidence interval, and display it in the figure.

Results

Identification of DEGs in TNBC

The Fig. 1 presented the workflow of this study. There were 76 TNBC tissues and 175 non-TNBC tissues in the current research. 147 and 391 DEGs from GSE36693 and GSE65216 were respectively extracted by GEO2R online tools. 157 up-regulated and 130 down-regulated DEGs were conducted in GSE36693 (Fig. 2A) with the criteria of |logFC| > 2 and adjusted p-value < 0.05; 282 up-regulated and 249 down-regulated DEGs were gained in GSE65216 (Fig. 2B). Then, Venn diagram online tool was utilized to intersect the DEGs in the two datasets and visualize them to identify the common DEGs. Consequences were that totally 140 commonly DEGs were detected in the TNBC tissues, including 69 down-regulated genes with the limitation of logFC<-2, and 71 up-regulated genes with the limitation of logFC > 2 (Fig. 2C).

DEGs GO and KEGG pathway analysis in TNBC

All of the 140 DEGs were included by DAVID online tool and the results of GO analysis indicated that: 1) as for BP, those up-regulated genes were mainly enriched in peripheral nervous system development, epidermis development, single organismal cell-cell adhesion, positive regulation of transcription from RNA polymerase II promoter, cytoskeleton organization etc., while down-regulated DEGs were enriched in phosphatidylinositol 3-kinase signaling, positive regulation of transcription from RNA polymerase II promoter, regulation of intracellular transport, wound healing and so on; 2) in the aspect of MF, up-regulated DEGs were enriched in transcription factor activity, RNA polymerase II distal enhancer sequence-specific binding, chitinase activity, chitin binding, and down-regulated DEGs were significantly enriched in RNA polymerase II transcription factor binding, calcium ion binding, estrogen response element binding, dystroglycan binding, transcription regulatory region DNA binding; 3) for CC, up-regulated DEGs were notably enriched in the extracellular space, extracellular exosome, epidermal lamellar body, intermediate filament while down-regulated genes were mainly enriched in extracellular space (Table 1).

Table 1. Gene Ontology analysis of differentially expressed genes in TNBC.

Expression	Category	Term	Count	%	P Value
Up-regulated	GOTERM_BP_DIRECT	GO:0007422~peripheral nervous system development	3	4.35	0.0036254
	GOTERM_BP_DIRECT	GO:0008544~epidermis development	4	5.80	0.003993946
	GOTERM_BP_DIRECT	GO:0016337~single organismal cell-cell adhesion	4	5.80	0.006458971
	GOTERM_CC_DIRECT	GO:0005615~extracellular space	16	23.19	7.34E-05
	GOTERM_CC_DIRECT	GO:0005882~intermediate filament	4	5.80	0.008000418
	GOTERM_CC_DIRECT	GO:0070062~extracellular exosome	19	27.54	0.009474606
	GOTERM_CC_DIRECT	GO:0097209~epidermal lamellar body	2	2.90	0.014409065
	GOTERM_MF_DIRECT	GO:0004568~chitinase activity	2	2.90	0.024620645
	GOTERM_MF_DIRECT	GO:0008061~chitin binding	2	2.90	0.028088866
Down -regulated	GOTERM_BP_DIRECT	GO:0014065~phosphatidylinositol 3-kinase signaling	3	0.03	0.004160756
	GOTERM_BP_DIRECT	GO:0045944~positive regulation of transcription from RNA polymerase II promoter	10	0.10	0.00769136
	GOTERM_CC_DIRECT	GO:0005615~extracellular space	12	0.12	0.007648324
	GOTERM_MF_DIRECT	GO:0001085~RNA polymerase II transcription factor binding	3	0.03	0.012119346
	GOTERM_MF_DIRECT	GO:0005509~calcium ion binding	8	0.08	0.013256787
	GOTERM_MF_DIRECT	GO:0034056~estrogen response element binding	2	0.02	0.014142799

Outcome of KEGG analysis were shown in Table 2 and Fig. 3B, indicating that DEGs were significantly enriched in p53 signaling pathway, prostate cancer and metabolic pathways (P<0.05).

Table 2. KEGG pathway analysis of differentially expressed genes in TNBC.

Category	Term	Count	%	P Value
KEGG_PATHWAY	hsa04115: p53 signaling pathway	4	2.88	0.008797166
KEGG_PATHWAY	hsa05215: Prostate cancer	4	2.88	0.018367061
KEGG_PATHWAY	hsa01100: Metabolic pathways	14	10.07	0.037159241

Protein–protein interaction network (PPI) analysis

The 140 DEGs including 69 up-regulated genes and 71 down-regulated genes were imported into the DEGs PPI network complex which contained 94 nodes and 180 edges (Fig. 4A). And 46 isolated nodes were excluded. Then, Cytoscape MCODE was applied and 29 central nodes were identified of the 94 nodes (Fig. 4B).

Re-analysis of 29 selected genes by GO and KEGG pathway enrichment

For the sake of figuring out the possible pathways of the 29 selected DEGs and verify whether they were consistent with the result of 140 DEGs enrichment analysis, the GO and KEGG pathway enrichment analysis was performed on these 29 DEGs again by DAVID (P<0.05). Outcome revealed that the 29 selected DEGs evidently enriched in the p53 signaling pathway, pathways in cancer, oocyte meiosis and prostate cancer (P<0.05, Tables 3 and 4 & Fig. 5).

Table 3. Re-analysis of 29 selected genes via Gene Ontology enrichment.

Category	Term	Count	P Value
GOTERM_BP_DIRECT	GO:0045893~positive regulation of transcription, DNA-templated	9	9.84E-07
GOTERM_BP_DIRECT	GO:0045944~positive regulation of transcription from RNA polymerase II promoter	11	1.50E-06
GOTERM_BP_DIRECT	GO:0006366~transcription from RNA polymerase II promoter	7	1.35E-04
GOTERM_BP_DIRECT	GO:0048546~digestive tract morphogenesis	3	1.92E-04
GOTERM_BP_DIRECT	GO:0043568~positive regulation of insulin-like growth factor receptor signaling pathway	3	1.92E-04
GOTERM_BP_DIRECT	GO:0014065~phosphatidylinositol 3-kinase signaling	3	8.52E-04
GOTERM_CC_DIRECT	GO:0005882~intermediate filament	4	6.79E-04
GOTERM_CC_DIRECT	GO:0005615~extracellular space	9	7.11E-04
GOTERM_CC_DIRECT	GO:0005634~nucleus	16	0.006994347
GOTERM_CC_DIRECT	GO:0070062~extracellular exosome	10	0.020917677
GOTERM_CC_DIRECT	GO:0000790~nuclear chromatin	3	0.035216711
GOTERM_CC_DIRECT	GO:0030057~desmosome	2	0.036252678
GOTERM_MF_DIRECT	GO: 0008134~transcription factor binding	7	5.93E-06
GOTERM_MF_DIRECT	GO: 0043565~sequence-specific DNA binding	7	1.72E-04
GOTERM_MF_DIRECT	GO: 0000981~RNA polymerase II transcription factor activity, sequence-specific DNA binding	5	1.72E-04
GOTERM_MF_DIRECT	GO: 0044212~transcription regulatory region DNA binding	5	3.98E-04
GOTERM_MF_DIRECT	GO: 0005515~protein binding	24	9.40E-04
GOTERM_MF_DIRECT	GO: 0001085~RNA polymerase II transcription factor binding	3	0.002738639

Table 4 . Re-analysis of 29 selected genes via KEGG pathway enrichment.

Category	Term	Count	P Value	Genes
KEGG_PATHWAY	hsa04115: p53 signaling pathway	4	1.38E-04	CCNE1, CDKN2A, SERPINB5, IGF1
KEGG_PATHWAY	hsa05215: Prostate cancer	3	0.008256173	CCNE1, AR, IGF1
KEGG_PATHWAY	hsa04114: Oocyte meiosis	3	0.012905535	CCNE1, AR, IGF1
KEGG_PATHWAY	hsa05200: Pathways in cancer	4	0.021651652	CCNE1, AR, CDKN2A, IGF1

Analysis of core genes by the GEPIA

After taking the results of both PPI analysis and the KEGG pathway enrichment into consideration, we found that CCNE1, CDKN2A, AR, SERPINB5 and IGF1 among the 29 selected genes could play a key role in common significantly enriched pathways. Then, GEPIA was utilized to analyze the differences in the expression of these 5 genes of TNBC and non-TNBC tissues. Compared with non-TBNC samples, three hub-genes (CCNE1, AR, CDKN2A) were highly expressed while two gens (SERPINB5, IGF1) were not in TNBC samples ($P<0.05$, Fig. 6).

Survival analysis using cBioportal

cBioportal was used to identify 3 hub genes using survival data. Only CCNE1 had a obviously worse survival, while the other two genes did not ($P < 0.05$, Fig. 7).

Discussion

In order to recognize meaningful prognostic biomarkers of TNBC, this research analyzed two profile data sets (GSE36693 and GSE65216) using bioinformatics methods. A total of 76 TNBC specimens and 175 non-TNBC specimens were included in the current research. With analysis of GEO2R and Venn diagram, 140 common DEGs were identified ($|logFC|>2$ and adjust $P<0.01$). Through KEGG pathway analysis, DEGs were obviously enriched in p53 signaling pathway, prostate cancer and metabolic pathways ($P<0.05$). Then, by using STRING online tool as well as Cytoscape, we constructed DEGs PPI network complex composed of 94 nodes and 180 edges. In addition, 29 vital genes were screened from the PPI network complex via the analysis of Cytoscape's MCODE. Next, we performed KEGG enrichment analysis again on the 29 DEGs through DAVID and found that these genes were enriched in the p53 signaling pathway, pathways in cancer, oocyte meiosis and prostate cancer ($P<0.05$). Taking the results of both PPI analysis and the KEGG pathway enrichment into consideration, CCNE1, CDKN2A, AR, SERPINB5 and IGF1 of the 29 selected genes were found to play a key role in common significantly enriched pathways. Three hub-genes showed high expression with the validation of GEPIA expression sequencing. Finally, only CCNE1 had a significantly worse survival identified by survival analysis using cBioportal, which could be used as a new potential target for providing new treatment ideas for TNBC and improving prognosis.

CCNE1, also known as cyclin E1, is encoded by this gene belongs to the highly conserved cyclin family. The characteristic of its members is that the protein concentration changes drastically with the cycle throughout the cell cycle. Cyclin is a regulator of cyclin-dependent kinase (CDK), CCNE1 forms a complex with CDK2 as a regulatory subunit whose activity is necessary for cell cycle G1/S transition [21]. The protein is abundantly present at the boundary of G1-S phase and degraded as the cell cycle passes through S phase. This gene has been observed to be highly expressed in many tumors, which can lead to chromosomal instability and contribute to tumorigenesis. The dysregulation of CCNE1-CDK2 activity is related to a variety of cancers including nasopharyngeal carcinoma, bladder cancer and breast cancer, and has been fully proven [22-24]. Accumulating data proved that TNBC frequently expressed CCNE1, while ER positive cancer did not [25], and the absence of CCNE1 for poorer DFS [25].

In the p53 pathway, p53 acted as a tumor suppressor gene. Contrary to the activation of p53 regulatory checkpoint or apoptosis, the expression of cyclin E protein promotes the process of entering S phase from G1 phase. And the lack of p53 function gives tumor cells an escape gap, so that tumor cells can avoid cell cycle arrest or cell death, and advance to the next stage through this disorder and uncontrolled growth [21, 26]. In addition, the loss of functional expression of the G1 checkpoint CDK inhibitor, p21 is also related to the carcinogenesis and disease progression of breast cancer, at the same time, more and more data indicate that the loss of function of p21 can mediate the drug-resistant phenotype which always means a poor prognosis [27]. In our bioinformatics analysis, the phenomenon of the up-regulated CCNE1 enriched in the p53 pathway was verified again and it CCNE1 really plays an important role in TNBC, but relative clinical practice is lacking.

As for pathways in cancer, specifically in the cell cycle, the transcription factor E2F1 and the tumor suppressor protein retinoblastoma (RB) are two key factors that regulate the progression of the cell cycle. They determine whether the cell can carry out the process of DNA replication and cell division by regulating the checkpoints of G1/S

and G2/M together [28]. The CCNE1/CDK2 complex can phosphorylate RB, then release E2F1 and activate its transcriptional activity to advance the cell cycle from G1 to S phase, while dephosphorylation of RB promotes E2F1 heterodimerization while inhibiting E2F1 activity [29, 30]. It can be seen in the Fig. 7B that Cyclin E-CDK2 is associated with RB gene in the process of pathways in cancer, and this enrichment of CCNE1 was validated by us as well. There is no doubt that CCNE1 can be a useful target of TNBC in the future.

Numerous studies have proved that CCNE1 was related to the carcinogenesis and development of various types of cancer. Nevertheless, researches and clinical practice reported about this gene in TNBC is insufficient, in other words, they haven't been taken seriously enough. Further experiments should be carried out on CCNE1 to better validate its functions. Thus, results in our research may provide useful information and directions for prospective research on TNBC treatment and prognosis. But this needs to be further verified through experiments in vitro or in vivo.

Conclusions

CCNE1 could lead to a poorer prognostic in TNBC identified via bioinformatic analysis and plays a key role in the progression of TNBC which may contribute potential targets for the diagnosis, treatment and prognosis assessment of TNBC.

List Of Abbreviations

TNBC: Triple negative breast cancer

HER2: human epidermal growth factor receptor 2

DEGs: differentially expressed genes

HR: homologous recombination

GEO: Gene Expression Omnibus

GO: Gene Ontology

KEGG: Kyoto Encyclopedia of Gene and Genome

BP: biological process

CC: cellular component

MF: molecular function

PPI: protein-protein interaction

MCODE: Molecular Complex Detection

GEPIA: Gene Expression Profiling Interactive Analysis

DFS: overall survival (OS) and distant free survival

CDK: cyclin-dependent kinase

Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: All the data in this research are available from GEO database and TCGA database.

Competing interests: The authors declare that they have no competing interests.

Funding: None.

Authors' contributions:

Gaosong Wu and Qianqian Yuan was involved in data methodology

Qianqian Yuan and Yiqin Liao were involved in data curation, formal analysis, and methodology.

Gaosong Wu were involved in supervision.

Qianqian Yuan and Lewei Zheng

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018. https://doi.org/10.3322/caac.21492
2. Siegel RL, Miller KD, Jemal A: Cancer statistics, 2019. CA Cancer J Clin 2019. https://doi.org/10.3322/caac.21551
3. Rakha EA, Reis-Filho JS, Ellis IO: Basal-like breast cancer: a critical review. J Clin Oncol 2008. https://doi.org/10.1200/JCO.2007.13.1748
4. Vagia E, Mahalingam D, Cristofanilli M: The Landscape of Targeted Therapies in TNBC. Cancers (Basel) 2020. https://doi.org/10.3390/cancers12040916
5. Malorni L, Shetty PB, De Angelis C, Hilsenbeck S, Rimawi MF, Elledge R,... Arpino G: Clinical and biologic features of triple-negative breast cancers in a large cohort of patients with long-term follow-up. Breast Cancer Res Treat 2012. https://doi.org/10.1007/s10549-012-2315-y
6. Judes G, Rifaï K, Daures M, Dubois L, Bignon YJ, Penault-Llorca F,... Bernard-Gallon D: High-throughput «Omics» technologies: New tools for the study of triple-negative breast cancer. Cancer Lett 2016. https://doi.org/10.1016/j.canlet.2016.03.001
7. Evans DG, Howell A, Ward D, Lalloo F, Jones JL, Eccles DM: Prevalence of BRCA1 and BRCA2 mutations in triple negative breast cancer. J Med Genet 2011. https://doi.org/10.1136/jmedgenet-2011-100006
8. Roy PG, Thompson AM: Cyclin D1 and breast cancer. Breast 2006. https://doi.org/10.1016/j.breast.2006.02.005

9. Edgar R, Domrachev M, Lash AE: Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002. https://doi.org/10.1093/nar/30.1.207
10. Davis S, Meltzer PS: GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics 2007. https://doi.org/10.1093/bioinformatics/btm254
11. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM,... Sherlock G: Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000. https://doi.org/10.1038/75556
12. Kanehisa M, Goto S: KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000. https://doi.org/10.1093/nar/28.1.27
13. Huang DW, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009. https://doi.org/10.1038/nprot.2008.211
14. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J,... von Mering C: STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 2015. https://doi.org/10.1093/nar/gku1003
15. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D,... Ideker T: Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003. https://doi.org/10.1101/gr.1239303
16. Bandettini WP, Kellman P, Mancini C, Booker OJ, Vasu S, Leung SW,... Arai AE: MultiContrast Delayed Enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study. J Cardiovasc Magn Reson 2012. https://doi.org/10.1186/1532-429X-14-83
17. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z: GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017. https://doi.org/10.1093/nar/gkx247
18. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO,... Schultz N: Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013. https://doi.org/10.1126/scisignal.2004088
19. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA,... Schultz N: The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012. https://doi.org/10.1158/2159-8290.CD-12-0095
20. Zheng H, Zhang G, Zhang L, Wang Q, Li H, Han Y,... Guo X: Comprehensive Review of Web Servers and Bioinformatics Tools for Cancer Prognosis Analysis. Front Oncol 2020. https://doi.org/10.3389/fonc.2020.00068
21. Lopez-Beltran A, MacLennan GT, Montironi R: Cyclin E as molecular marker in the management of breast cancer: a review. Anal Quant Cytol Histol 2006.
22. Chen M, Wu R, Li G, Liu C, Tan L, Xiao K,... Qin Z: Motor neuron and pancreas homeobox 1/HLXB9 promotes sustained proliferation in bladder cancer by upregulating CCNE1/2. J Exp Clin Cancer Res 2018. https://doi.org/10.1186/s13046-018-0829-9
23. Wu J, Lv Q, He J, Zhang H, Mei X, Cui K,... Zhang Y: MicroRNA-188 suppresses G1/S transition by targeting multiple cyclin/CDK complexes. Cell Commun Signal 2014. https://doi.org/10.1186/s12964-014-0066-6
24. Luo Q, Li X, Li J, Kong X, Zhang J, Chen L,... Fang L: MiR-15a is underexpressed and inhibits the cell cycle by targeting CCNE1 in breast cancer. Int J Oncol 2013. https://doi.org/10.3892/ijo.2013.2034

25. Zagouri F, Kotoula V, Kouvatseas G, Sotiropoulou M, Koletsa T, Gavressea T,... Fountzilas G: Protein expression patterns of cell cycle regulators in operable breast cancer. Plos One 2017. https://doi.org/10.1371/journal.pone.0180489
26. Barton MC, Akli S, Keyomarsi K: Deregulation of cyclin E meets dysfunction in p53: closing the escape hatch on breast cancer. J Cell Physiol 2006. https://doi.org/10.1002/jcp.20818
27. Abukhdeir AM, Park BH: P21 and p27: roles in carcinogenesis and drug resistance. Expert Rev Mol Med 2008. https://doi.org/10.1017/S1462399408000744
28. Kanska J, Zakhour M, Taylor-Harding B, Karlan BY, Wiedemeyer WR: Cyclin E as a potential therapeutic target in high grade serous ovarian cancer. Gynecol Oncol 2016. https://doi.org/10.1016/j.ygyno.2016.07.111
29. Sheldon LA: Inhibition of E2F1 activity and cell cycle progression by arsenic via retinoblastoma protein. Cell Cycle 2017. https://doi.org/10.1080/15384101.2017.1338221
30. Liu F, Cai Y, Rong X, Chen J, Zheng D, Chen L,... Ruan J: MiR-661 promotes tumor invasion and metastasis by directly inhibiting RB1 in non-small cell lung cancer. Mol Cancer 2017. https://doi.org/10.1186/s12943-017-0698-4

Figures

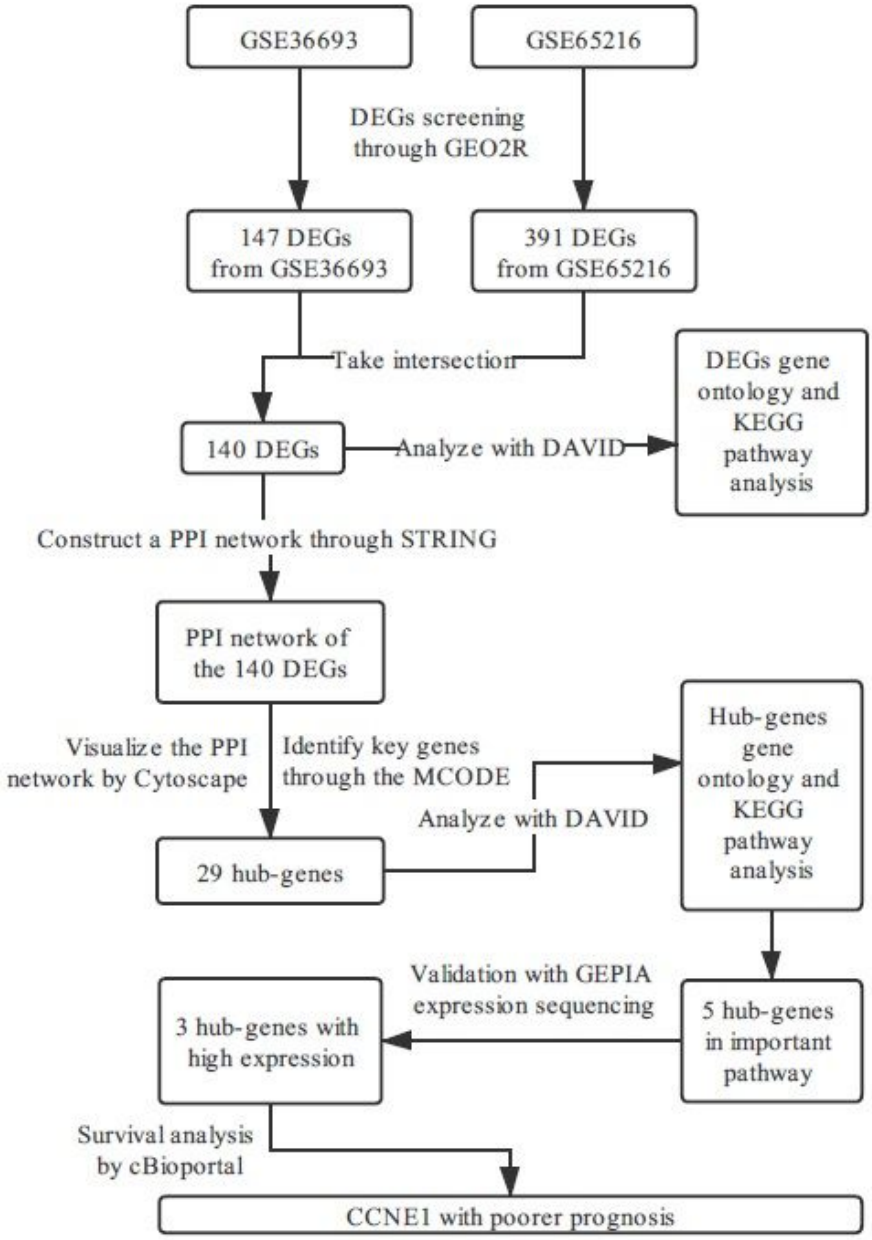

Figure 1

Flow diagram of the data selection.

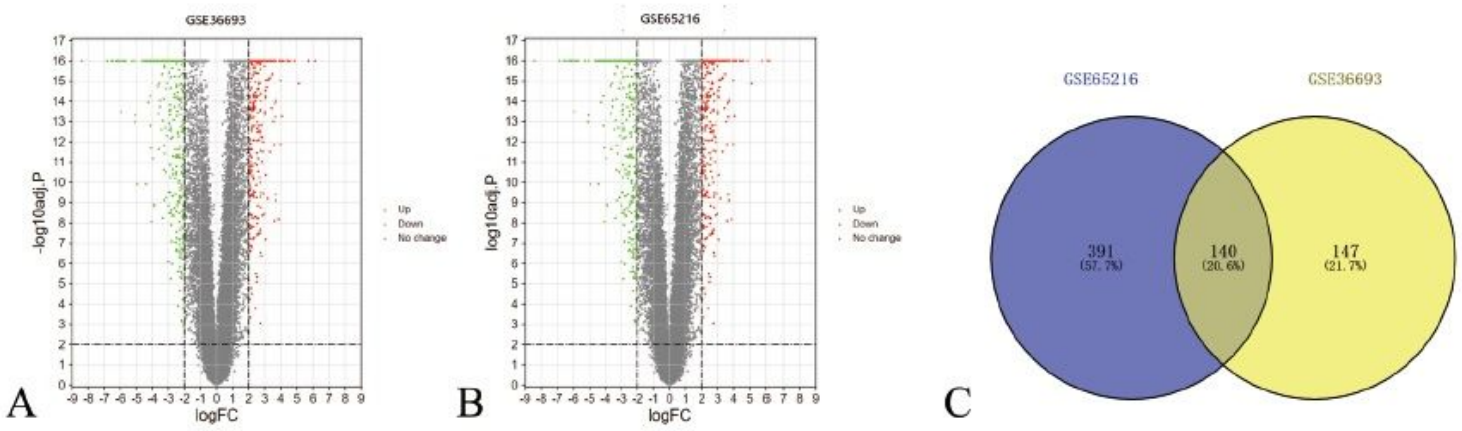

Figure 2

Volcano plots for DEGs in TNBC and non-TNBC tissues based on data from the GEO datasets. (A) GSE36693, (B) GSE65216. (C) Take the intersection of the DEGs in the two data sets (GSE36693 and GSE65216) via the Venn diagram online tool, and get 140 common genes. Different colors in the figure mean different data sets.

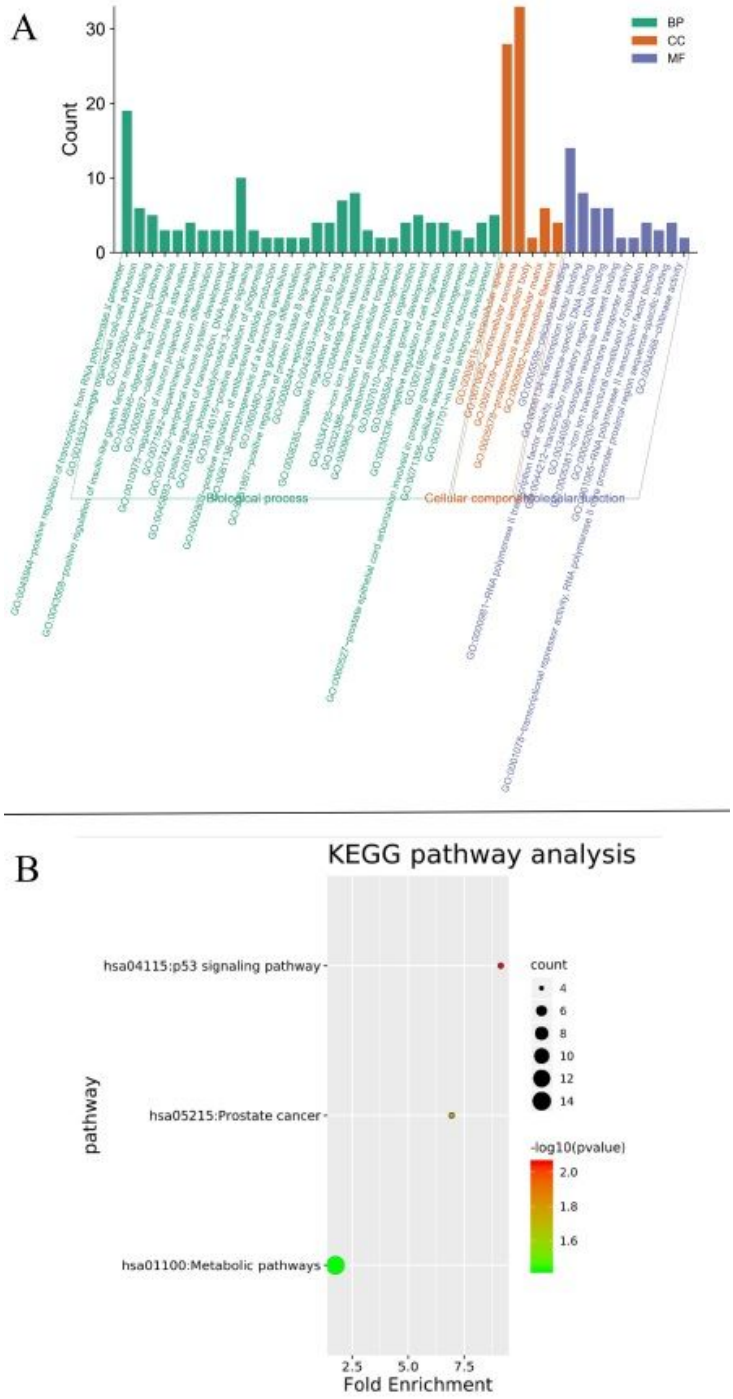

Figure 3

DAVID analysis of DEGs. (A) GO functional annotation of top 45 enrichment terms. (B) KEGG pathway enrichment analysis of top three enrichment terms. The count of genes enriched in terms is indicated by the node size; the P value is shown by the color, the redder the color, the more significant it is. DAVID, Database for Annotation, Visualization, and Integrated Discovery; DEGs, differentially expressed genes; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

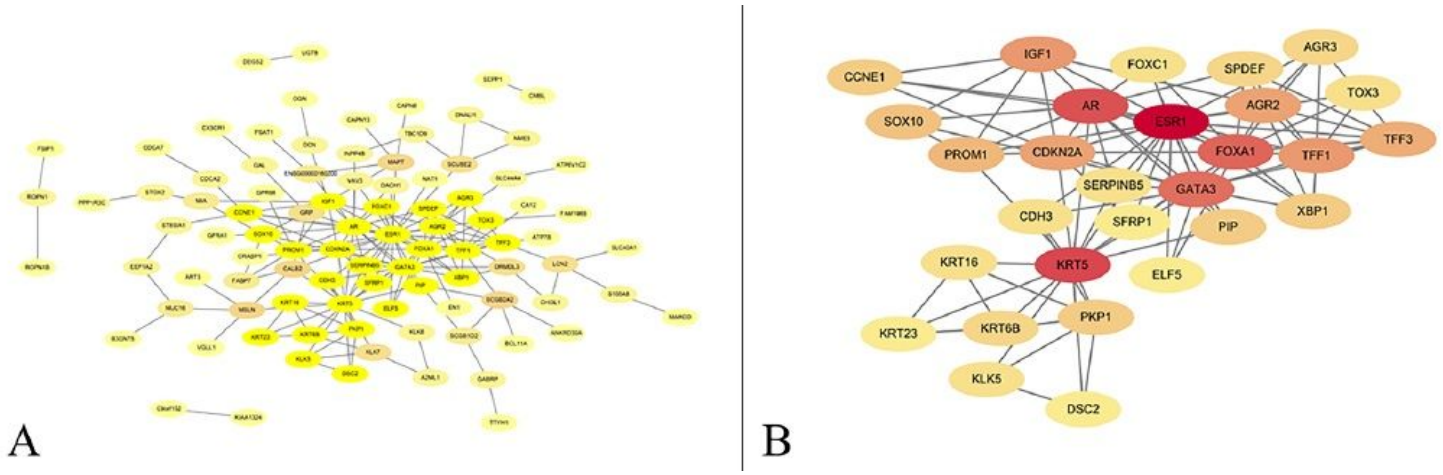

Figure 4

Common DEGs PPI network set up by STRING database and Module analysis. (A) A total of 140 DEGs were used to construct a DEG PPI network complex, and 46 isolated nodes were removed. The nodes meant proteins; the edges meant the interaction of proteins. (B) Module analysis via Cytoscape software selected 29 hub-genes (degree cut-off=2, node score cut-off=0.2, k-core= 2, and max. Depth= 100).

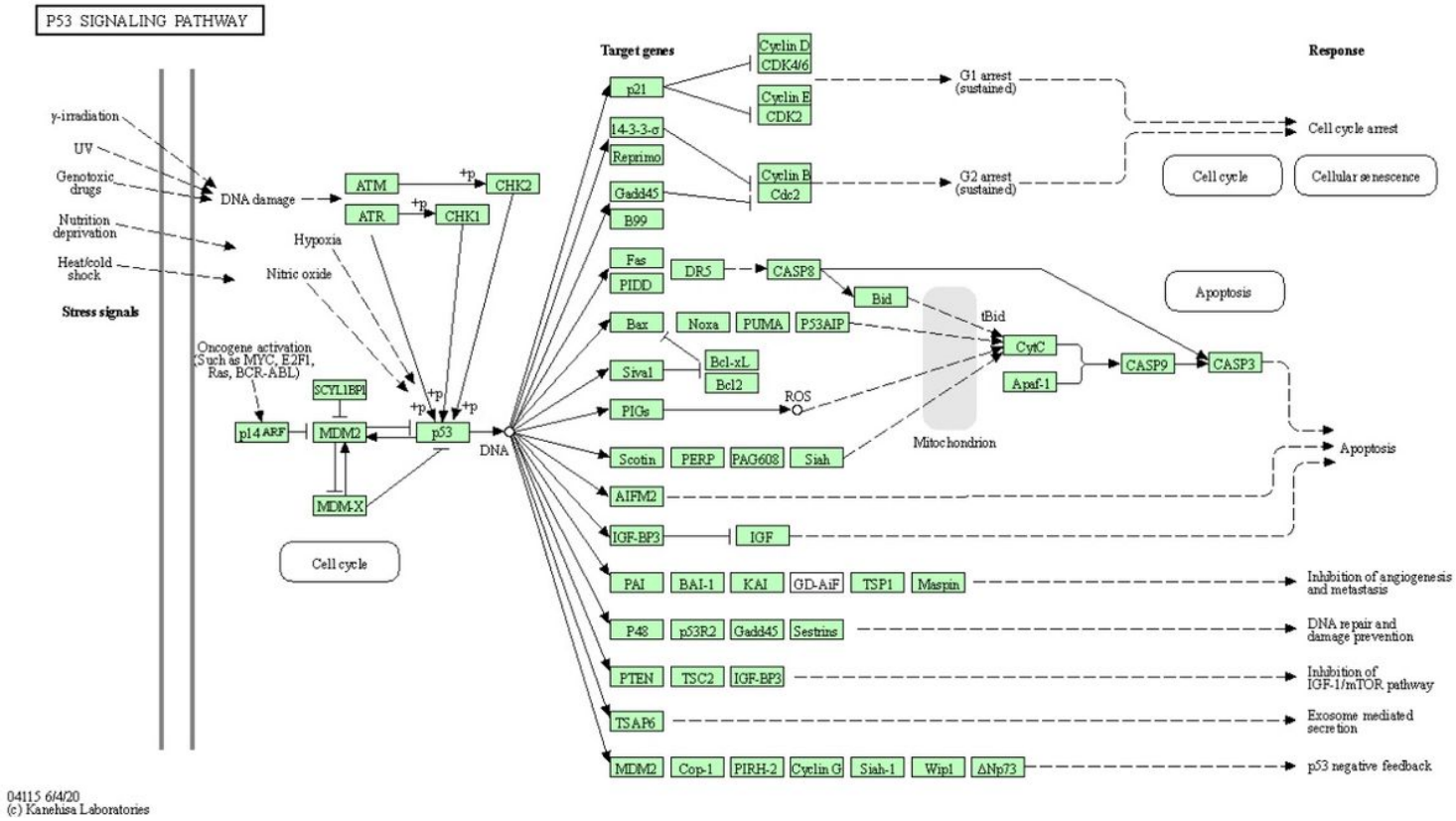

Figure 5

Re-analysis of 29 selected genes by KEGG pathway enrichment. (A) Schematic diagram of p53 signaling pathway. (B) Schematic diagram of pathways in cancer. Cyclin E means CCNE1. CycB means CCNB1.

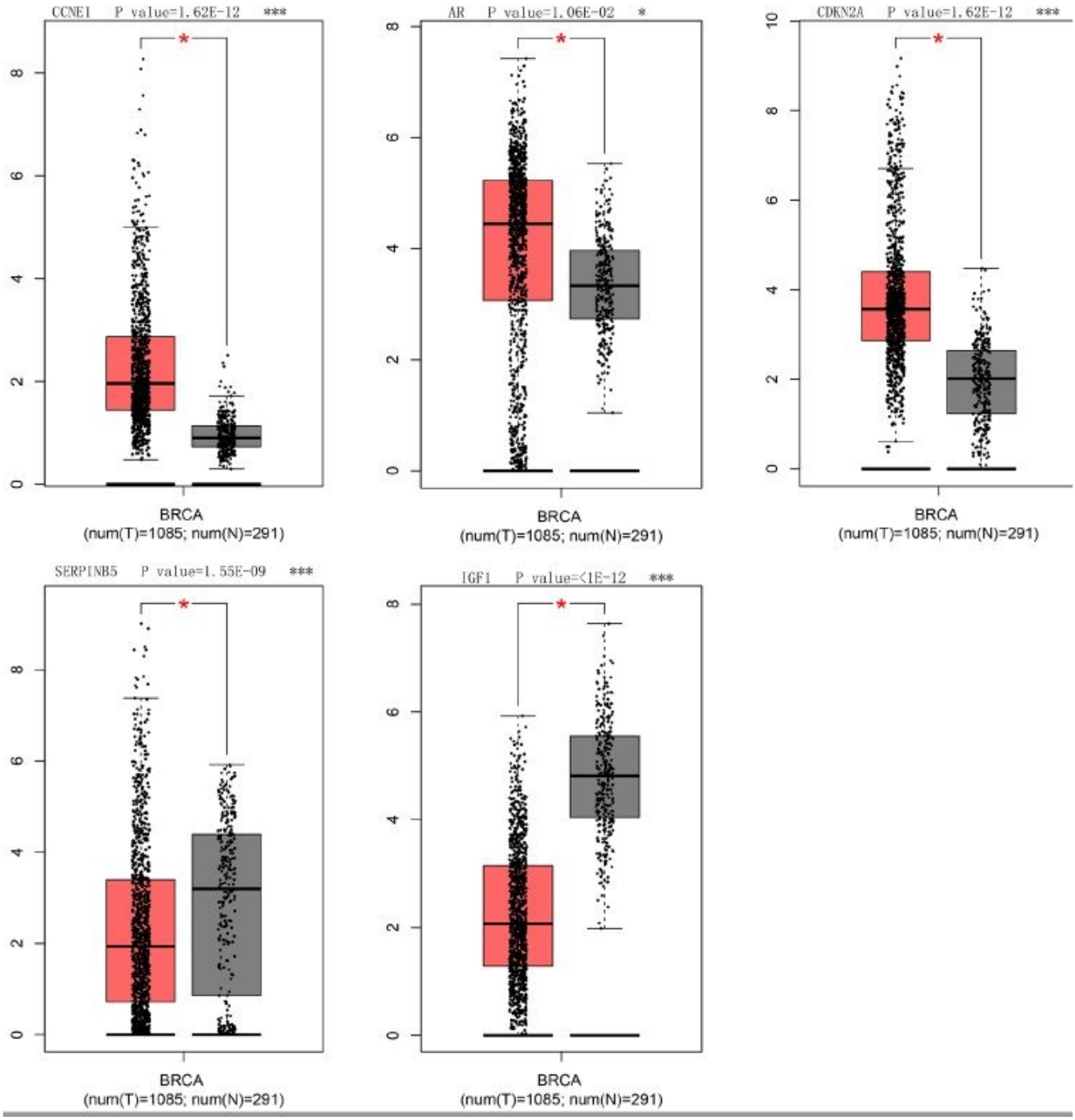

Figure 6

Significantly expressed five genes in TNBC samples compared to normal samples. Three of five genes had notable high expression in TNBC specimen compared to normal specimen (*P<0.05). Red color refers to tumor tissues and blue color refers to normal samples.

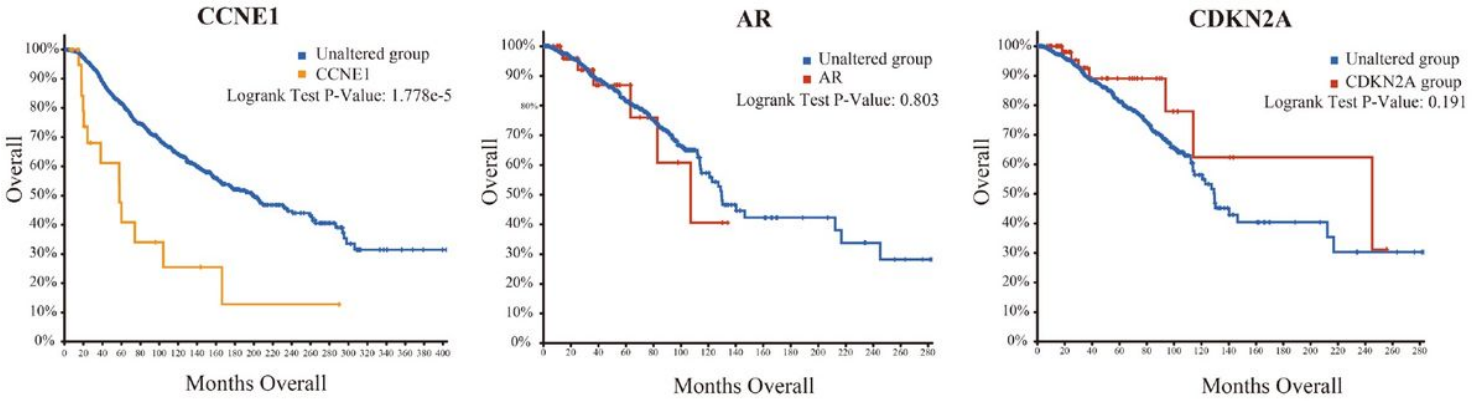

Figure 7

The prognostic information of the three hub-genes. cBioportal were utilized to identify the prognositc information of the three core genes, CCNE1 had a significantly worse survival rate (P< 0.05).