Integrated bioinformatics analysis identifies core genes in breast cancer

Niu Yuqin

The First Affiliated Hospital of school of Medicine

Zhang Xue

Shihizi University School of Medicine

Wei Yu

The First Affiliated Hospital of school of Medicine

Zhang Shuai

Shihezi University School of Medicine

Song Yin

Shihezi University School of Medicine

Sun Jiawei

Shihezi University School of Medicine

Li Yang (✉ 1443468586@qq.com)

Department of Biochemistry of Molecular Biology, Tongji Medical college

Zhang Zhiwei

Shihezi University School of Medicine

Research article

Keywords: Breast cancer, bioinformatics analysis, biomarker, therapeutic

Posted Date: June 3rd, 2019

DOI: <https://doi.org/10.21203/rs.2.10014/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background Breast cancer (BRCA) is one of the most common malignancies in women. To reveal the molecular mechanism of the BRCA, the core genes associated with BRCA was studied by the integrated bioinformatics methods in the current study. Results A total of 1181 DEGs, 84 DEGs, and 190 DEGs were detected in GSE7904, GSE10797, and GSE103512, respectively. In addition, 7 hub genes overlapped from 3 datasets were selected, and ANXA1, MYH11, IGFBP6, FOS, FOSB, KIT, and WLS might be the core genes of BRCA. The Go function analysis showed that the DEGs were enriched in chromosome segregation, gland development, leukocyte migration, extracellular structure organization and so on. The KEGG pathway analysis showed that the most DEGs were enriched in Pathways in cancer, Focal adhesion, ECM-receptor interaction, PPAR signaling pathway, etc. Additionally, ANXA1 was associated with worse overall survival, and the other 6 hub genes were associated with better overall survival. Conclusions ANXA1, MYH11, IGFBP6, FOS, FOSB, KIT, and WLS might be the core genes of BRCA. WLS, IGFBP6, and FOSB might be a new biomarker of the breast tumor and potential new drug targets. This study provided new evidence for the future genomic individualized treatment of BRCA.

Background

Benefits from early detection and treatment improvements, the survival rate of cancer increased in recent years. However, it still poses a great threat to human health and is a major public health problem. Breast cancer (BRCA) is the first cause of death of women in the world[1][1], and it also is a big public health problem in China [2, 3][2, 3].

Because of its poor prognosis and short average survival period, the early diagnosis and treatment of BRCA is very important for the survival. Previous reports showed that BRCA is highly heterogeneous [4][4], and the traditional anatomical stage and histological classification could no meet the clinical needs. Thus, Molecular level study of pathogenesis, treatment, and prognosis has become the current research hotspots in the field of BRCA [5, 6][5, 6].

In recent years, numerous biomarkers and therapeutic target genes for BRCA have been discovered, for example, estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor-2 (HER2); and most of them were useful in the diagnose and identify subtypes of BRCA [7][7]. Using biomarkers to diagnose cancer early is a good choice for the diagnose of BRCA [8][8].

Over the last decade, comprehensive sequencing efforts have revealed the genomic landscapes of common forms of human cancer [9]. With the help of high-through technology, a large number of data about BRCA, especially the data of cDNA microarray, can get and used in integrated informatics studies now. GEO (gene expression omnibus) is the largest and most comprehensive public gene expression database resource available today [9][9]. In this study, three datasets (GSE7904, GSE103512, GSE10797) were downloaded from GEO for analysis using R [10][10]. The screening of DEGs, the biological processes (BP) and KEGG pathway were analyzed with R packages. We used the GEPIA [11][11] database to observe changes in the expression levels of these genes between cancer and normal population. In addition, the correlation between genes, survival, and disease was tested based on Kaplan Meier plotter online database (<http://kmplot.com/analysis/>) [12][12]

Methods

Raw data of BRCA microarray from the GEO database was subjected to quantity control (QC). the differentially expressed genes (DEGs) were studied using the “limma” package of R by the condition of both adj. p-value < 0.05,

and log2 fold change of >1 or <-1 . Then DEGs were analyzed by GO function and KEGG. In addition, the core genes found were identified by the performance of expression level and survival analysis in BRCA.

Search strategy and data collection

Public microarray repositories were curated to search through the Gene Expression Omnibus database GEO (<http://www.ncbi.nlm.nih.gov/geo>) database using the keywords "BRCA", or "mammary cancer". The datasets were reviewed manually and only datasets which fulfill the following criteria were included for further analysis: (1) Expression profiling by array, (2) Studies which comprised CEL raw files, and (3) Paired normal-tumor breast tissue samples. 3 microarray datasets of GSE7904, GSE10797 and GSE103512 were collected.

Quality control of microarray

Quality control is mainly focused on CEL file-level processing, like the simplest visual observation, the averaging method and more advanced data fitting methods. The three levels of quality control functions are implemented by the 'image' function, the 'simpleaffy' package, and the 'affyPLM' package. After quality control, unqualified sample data were removed.

Individual microarray data analysis

R tools were used in this study. All datasets were normalized individually on the base-2 logarithm by Robust Multi-Array Average (RMA) and Linear Models for Microarray (LIMMA) package and annotated by converting different probe IDs to gene IDs.

Genes significantly dysregulated in BRCA as compared to matched data at normal samples were defined by adj. p-value < 0.05 , and log2 fold change of >1 (upregulated genes) or <-1 (downregulated genes).

DEGs enrichment analysis and its visualization

Gene enrichment analysis concluding GO enrichment analysis and KEGG enrichment analysis was done by the 'GStats' and 'GeneAnswers' package. We used the 'Clusterfiler' and 'pheatmap' package to visualizing the statistical analysis.

Survival analysis of key genes

The KM-PLOT is capable to assess the effect of 54,675 genes on survival using 10,461 cases of cancer samples, including 5,143 cases of breast cancers, 1,816 cases of ovarian cancers, 2,437 cases of lung cancers and 1,065 cases of gastric cancers with a mean follow-up of 69, 40, 49 and 33 months, respectively [19]. The primary purpose of the tool is a meta-analysis based on biomarker assessment [19]. In this study, key genes ("MYH11", "IGFBP6", "WLS", "ANXA1", "FOSB", "KIT", "FOS") were assessed for their effect on survival with with 95% confidence intervals and log rank P value <0.01 .

Correlation between hub genes and BRCA

Study on the correlation between hub genes and BRCA was performed by the online tools GEPIA (<http://gepia.cancer-pku.cn/index.html>).

Function protein association networks analysis

The function protein association networks of these three proteins through the “STRING” website (<https://string-db.org/cgi/input.pl>), and found that

Results

Quality control of microarray

Quality control of the three data sets of microarrays was conducted. The results showed that the overall chip quality of GSE7904 was good and RNA degradation was small (Figure.S1). The overall chip quality of GSE10797 is general, and RNA degradation was small (Figure.S2). and the GSE103512 chipset quality is slightly worse, RNA degradation is obvious (Figure.S3). based on these quality control analysis, 4 data of chips, 14 data of chips in the GSE10797 and GSE103512, respectively, were removed in the following studies (see supply information).

the normalization of the three datasets showed that between the experimental group and the control group, the expression of most genes detected were consistent, and the consistency between parallel experiments is stronger. Relative logarithmic expression (RLE) box plots reflect these trends showed that the overall trend of the fitting curve is consistent, indicating that the chip pretreatment effect is good (Figure 1). The normalized datasets were suitable for analysis of DEGs.

Significant Expression Changes of mRNAs in Breast Tumor

To determine the mRNA levels of genes in the breast tumor, we evaluated the mRNAs information in GSE7904, GSE10797, and GSE103512 by case-control design (Fig 2A, B, C). the results showed that contrast with the normal group, 338 genes were upregulated and 843 genes were downregulated in the GSE7904, 4 genes were upregulated and 80 genes were downregulated in GSE10797, and 48 genes were upregulated and 142 genes were downregulated in GSE103512, respectively (Figure 3A, B and C). We have listed the first ten DEGs sorted by the fold change (Table 1). Furthermore, levels of "MYH11", "IGFBP6", "WLS", "ANXA1", "FOSB", "KIT", "FOS" were different between tumor and normal tissues in the three data sets, indicating their potential involvement in the progression of breast cancer (Figure 2D). For further analysis, we chose the 7 genes evidently different mRNAs that were conserved among the three datasets: "MYH11", "IGFBP6", "WLS", "ANXA1", "FOSB", "KIT", "FOS". Cbioportal tools were used to check the 7 genes in the BRCA database (Fig.3D). The results showed the status of the 7 genes existed in BRCA and MYH11 was the significant one.

DEGs enrichment analysis and its visualization

GO and KEGG pathway analyses were performed to elucidate the biological function of the DEGs from three datasets (Fig .4). the results showed that in the data set of GSE7904, the significantly enriched GO terms at biological process domain: chromosome segregation, organelle fission, nuclear division, nuclear chromosome segregation, sister

chromatid segregation, mitotic nuclear division, mitotic sister chromatid segregation, regulation of mitotic nuclear division, regulation of chromosome separation, regulation of sister chromatid segregation. While the most obvious connection of pathways was, complement and coagulation cascades, ECM-receptor interaction, focal adhesion, Pathways in cancer, Cell cycle. In the data set of GSE10797, the significantly enriched GO terms at biological process domain were gland development, leukocyte migration, response to radiation, response to a steroid hormone, cognition, learning or memory, digestive system development, digestive tract development, cell-substrate junction assembly, response to progesterone. The pathways were Pathways in cancer, Focal adhesion, Hematopoietic cell lineage, Protein digestion and absorption, ECM-receptor interaction. In the data set of GSE103512, the enriched GO terms were extracellular structure organization, extracellular matrix organization, reactive oxygen species metabolic process, response to reactive oxygen species, cellular response to reactive oxygen species, glomerulus development, regulation of nitric oxide biosynthetic process, renal system vasculature development, glomerulus vasculature development, metanephric glomerulus development. The pathways were PPAR signaling pathway, Tyrosine metabolism, Arachidonic acid metabolism, Complement and coagulation cascades, and Cytokine-cytokine receptor interaction.

DEGs alterations may play important role in BRCA patient's survival

Survival analysis is widely used in clinical and epidemiological research. In this study, survival analysis was performed to assess the 7 genes in patient's survival (Fig 5). The results showed that the high expression of *FOS*, *FOSB*, *IGFBP6*, *KIT*, *MYH11* and *WLS* were associated with better overall survival (OS) for BRCA patients (HR>0.6, log rank P<0.0001), and the high expression of *ANXA1* was associated with worse overall survival (HR=1.2(1.06-1.35) log rank P= 0.0044, Figure 5). To determine the correlation between DEGs and BRCA, the expression levels of the 7 genes were evaluated by GEPIA, the result showed that the expression of these 7 genes was down-regulated in BRCA (P<0.05, Figure 6).

Various genes and medicines specific to DEGs reveal the 7 genes as potential targets for BRCA treatment

To evaluate if this DEGs could be targeted to BRCA therapeutics, a website integrated PhosphoSite, KEGG Drugs, pid, HumanCyc, Reactome, PANTHER, and DrugBank databases was used to analyze the correlation of these genes and drugs. Results showed that 22 kinds of FDA approved drugs were target to *ANXA1*, 11 kinds of FDA approved drugs and 10 kinds of non-FDA approved drugs were target to *KIT*, 2 kinds of FDA approved drugs were target to *MYH11*, and 1 kind of FDA approved, drugNadroparin, was target to *FOS* (Fig 7), however, there were no drug target to *IGFBP6*, *FOSB*, and *WLS* until now.

Discussion

BRCA is the most common cancer and the leading death cause among women younger than 45 years [13][13]. It is necessary to discover the molecular mechanism of the development of BRCA to decrease the mortality of BRCA.

The bioinformatics analysis is one of the useful methods for the study of molecular mechanism of cancer. In the current study, three GEO database, including GSE7904 (43 samples of BRCA and 19 normal samples), GSE10797 (65 samples of BRCA and 10 normal samples), and GSE103512 (65 samples of BRCA and 10 normal samples) were used to study the differential expression genes between BRCA and normal tissues. The results showed that a total of 1181 DEGs in GSE7904, 84 DEGs in GSE10797, and 190 DEGs in GSE103512 were found significantly expressed between BRCA and control group.

The GO analysis showed that the DEGs found in the dataset of GSE7904 were mainly involved in chromosome segregation, organelle fission, nuclear division. The DEGs found in the dataset of GSE7904 were mainly involved in gland development, leukocyte migration, response to radiation, response to a steroid hormone, cognition, learning or memory, digestive system development, and cell-substrate junction assembly. The DEGs found in the dataset of GSE103512 were mainly involved in extracellular structure organization, reactive oxygen species metabolic process, glomerulus development, and regulation of nitric oxide biosynthetic process. There was no common GO term among three datasets, indicated that the BRCA was highly of heterogeneity, consistent with the previous study [15].

Additionally, KEGG analysis showed that the DEGs found in the dataset of GSE7904 most connection of pathways included Complement and coagulation cascades, ECM-receptor interaction, Focal adhesion, Pathways in cancer, and Cell cycle. The DEGs found in the dataset of GSE7904 most obvious connection of pathways included Pathways in cancer, Focal adhesion, Hematopoietic cell lineage, Protein digestion and absorption, and ECM-receptor interaction. The DEGs found in the dataset of GSE103512 hub pathways were PPAR signaling pathway, Tyrosine metabolism, Arachidonic acid metabolism, Complement and coagulation cascades, and Cytokine-cytokine receptor interaction. There was no same pathways among three datasets, consistent with the GO analysis and previous report[15].

Based on the Venn diagram analysis, 7 hub genes were overlapped among 3 datasets, and they were ANXA1, MYH11, IGFBP6, FOS, FOSB, KIT, and WLS. In addition, the association of the seven hub genes with the overall survival of BRCA patients were studied, the results showed that ANXA1 was associated with worse overall survival, and the other 6 hub genes were associated with better overall survival. In addition, ANXA1, KIT, MYH11, and FOS have been targeted for drugs. There were no FDA drugs for IGFBP6, FOSB, and WLS. Further analysis showed that these three proteins interact with the well-known BRCA related proteins IGF1[14-16][14-16], FOS [17-19][17-19], and WNT1[20-22][20-22], suggested that IGFBP6, FOSB, and FOS might be useful target for BRCA treatment.

Conclusions

In summary, ANXA1, MYH11, IGFBP6, FOS, FOSB, KIT, and WLS were differentially expressed in breast tissue between BRCA and normal tissue, and maybe useful target for treatment and diagnosis of BRCA

Abbreviations

BRCA: Breast cancer KEGG: Kyoto Encyclopedia of Genes and Genomes .ER: Estrogen receptor ER. PR: Progesterone receptor. HER2: human epidermal growth factor receptor-2. GEO: gene expression omnibus. QC: quantity control. DEGs: the differentially expressed genes. RMA: Robust Multi-Array Average. LIMMA: Linear Models for Microarray.

Declarations

Acknowledgements

Not applicable.

Availability of data and materials

Not applicable.

Authors' contributions

The present review was mainly written by ZS and YL. All authors contributed toward data analysis, drafting and critically revising the paper, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note Springer

Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

Not applicable

References

1. DeSantis CE, Lin CC, Mariotto AB, Siegel RL, Stein KD, Kramer JL, Alteri R, Robbins AS, Jemal A: **Cancer treatment and survivorship statistics, 2014.** *CA Cancer J Clin* 2014, **64**(4):252-271.
2. Chu J, Zhou C, Guo X, Sun J, Xue F, Zhang J, Lu Z, Fu Z, Xu A: **Female Breast Cancer Mortality Clusters in Shandong Province, China: A Spatial Analysis.** *Sci Rep* 2017, **7**(1):105.

3. Lu S, Qian Y, Huang X, Yu H, Yang J, Han R, Su J, Du W, Zhou J, Dong M *et al*: **The association of dietary pattern and breast cancer in Jiangsu, China: A population-based case-control study.** *PLoS One* 2017, **12**(9):e0184453.
4. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, Watson M, Davies S, Bernard PS, Parker JS *et al*: **Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer.** *J Natl Cancer Inst* 2009, **101**(10):736-750.
5. Lindenberg MA, Miquel-Cases A, Retel VP, Sonke GS, Wesseling J, Stokkel MPM, van Harten WH: **Imaging performance in guiding response to neoadjuvant therapy according to breast cancer subtypes: A systematic literature review.** *Crit Rev Oncol Hematol* 2017, **112**:198-207.
6. Chae YK, Arya A, Chiec L, Shah H, Rosenberg A, Patel S, Raparia K, Choi J, Wainwright DA, Villafior V *et al*: **Challenges and future of biomarker tests in the era of precision oncology: Can we rely on immunohistochemistry (IHC) or fluorescence in situ hybridization (FISH) to select the optimal patients for matched therapy?** *Oncotarget* 2017, **8**(59):100863-100898.
7. Donepudi MS, Kondapalli K, Amos SJ, Venkanteshan P: **Breast cancer statistics and markers.** *J Cancer Res Ther* 2014, **10**(3):506-511.
8. Kulasingam V, Diamandis EP: **Strategies for discovering novel cancer biomarkers through utilization of emerging technologies.** *Nat Clin Pract Oncol* 2008, **5**(10):588-599.
9. Barrett T, Edgar R: **Mining microarray data at NCBI's Gene Expression Omnibus (GEO)*.** *Methods Mol Biol* 2006, **338**:175-190.
10. Davis S, Meltzer PS: **GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor.** *Bioinformatics* 2007, **23**(14):1846-1847.
11. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z: **GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses.** *Nucleic Acids Res* 2017, **45**(W1):W98-W102.
12. Szasz AM, Lanczky A, Nagy A, Forster S, Hark K, Green JE, Boussioutas A, Busuttill R, Szabo A, Gyorffy B: **Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients.** *Oncotarget* 2016, **7**(31):49322-49333.
13. Santos Sda S, Melo LR, Koifman RJ, Koifman S: **Breast cancer incidence and mortality in women under 50 years of age in Brazil.** *Cad Saude Publica* 2013, **29**(11):2230-2240.
14. Chen Z, Liu J, Chu D, Shan Y, Ma G, Zhang H, Zhang XD, Wang P, Chen Q, Deng C *et al*: **A dual-specific IGF-I/II human engineered antibody domain inhibits IGF signaling in breast cancer cells.** *Int J Biol Sci* 2018, **14**(7):799-806.
15. Daubriac J, Han S, Grahovac J, Smith E, Hosein A, Buchanan M, Basik M, Boucher Y: **The crosstalk between breast carcinoma-associated fibroblasts and cancer cells promotes RhoA-dependent invasion via IGF-1 and PAI-1.** *Oncotarget* 2018, **9**(12):10375-10387.
16. Belfiore A, Malaguarnera R, Nicolosi ML, Lappano R, Ragusa M, Morrione A, Vella V: **A novel functional crosstalk between DDR1 and the IGF axis and its relevance for breast cancer.** *Cell Adh Migr* 2018:1-10.
17. Fisler DA, Sikaria D, Yavorski JM, Tu YN, Blanck G: **Elucidating feed-forward apoptosis signatures in breast cancer datasets: Higher FOS expression associated with a better outcome.** *Oncol Lett* 2018, **16**(2):2757-2763.
18. Zhang Y, Zhao Y, Li H, Li Y, Cai X, Shen Y, Shi H, Li L, Liu Q, Zhang X *et al*: **The nuclear import of oncoprotein hepatitis B X-interacting protein depends on interacting with c-Fos and phosphorylation of both proteins in breast cancer cells.** *J Biol Chem* 2013, **288**(26):18961-18974.
19. Zhang X, Gamble MJ, Stadler S, Cherrington BD, Causey CP, Thompson PR, Roberson MS, Kraus WL, Coonrod SA: **Genome-wide analysis reveals PADI4 cooperates with Elk-1 to activate c-Fos expression in breast cancer**

cells. *PLoS Genet* 2011, **7**(6):e1002112.

20. Yan D, Avtanski D, Saxena NK, Sharma D: **Leptin-induced epithelial-mesenchymal transition in breast cancer cells requires beta-catenin activation via Akt/GSK3- and MTA1/Wnt1 protein-dependent pathways.** *J Biol Chem* 2012, **287**(11):8598-8612.
21. Benad P, Rauner M, Rachner TD, Hofbauer LC: **The anti-progestin RU-486 inhibits viability of MCF-7 breast cancer cells by suppressing WNT1.** *Cancer Lett* 2011, **312**(1):101-108.
22. Chiang KC, Yeh CN, Chung LC, Feng TH, Sun CC, Chen MF, Jan YY, Yeh TS, Chen SC, Juang HH: **WNT-1 inducible signaling pathway protein-1 enhances growth and tumorigenesis in human breast cancer.** *Sci Rep* 2015, **5**:8686.

Tables

Change	GSE7904			GSE10797			GSE103512		
	Gene symbol	Fold change	adj.P.Value	Gene symbol	Fold change	adj.P.Value	Gene symbol	Fold change	adj.P.Value
UP	NUSAP1	8.942491	5.77E-12	FN1	4.087731	3.81E-05	ESRP1	1.727936	1.21E-05
	CKS2	8.708405	1.31E-11	COL1A2	3.218053	0.000541	DNAH14	1.288937	1.34E-05
	RRM2	8.54353	2.34E-11	COL1A1	3.000341	0.001739	GRHL2	1.272979	3.65E-05
	DTL	8.513751	2.60E-11	COL3A1	2.23606	0.002082	COL10A1	2.520834	3.84E-05
	RRM2	8.368644	4.34E-11				FAM83H	1.084298	4.03E-05
	AURKA	8.260148	6.38E-11				STRBP	1.138334	9.90E-05
	CCNB2	8.240748	6.83E-11				EPS8L2	1.100007	0.000105
	KPNA2	8.196278	8.00E-11				HIST1H2BD	1.145106	0.000109
	ASPM	8.089972	1.17E-10				DDR1	1.246143	0.000111
	NUSAP1	8.942491	5.77E-12				ELF3	1.260238	0.000117
Down	OXTR	-4.3572	8.25E-22	KIT	-3.48393	1.01E-11	LYVE1	-2.82081	2.31E-09
	SAMD5	-3.1404	8.25E-22	CX3CL1	-2.57149	1.11E-09	ADAMTS5	-1.65617	3.28E-09
	TSHZ2	-2.82761	7.80E-20	NDRG2	-2.78323	2.21E-09	SCARA5	-1.4326	1.06E-08
	MYH11	-2.92615	1.18E-19	MFGE8	-2.22966	8.95E-09	ABCA8	-1.64912	4.47E-08
	DST	-3.20221	1.71E-19	NTRK2	-3.35141	1.35E-08	LILRB5	-1.06319	9.27E-08
	MAMDC2	-3.4483	4.82E-19	CX3CL1	-2.11445	3.13E-08	RHOJ	-1.13013	1.95E-07
	AK5	-3.41855	2.49E-18	SLC6A14	-3.22829	2.15E-07	TTN	-1.30037	3.05E-07
	HAS3	-2.68188	4.42E-17	DST	-4.49491	1.20E-06	BTNL9	-1.68343	3.50E-07
	COL17A1	-2.35717	2.41E-16	JUN	-2.17939	1.50E-06	IGFBP6	-1.49188	3.61E-07
	KLHL29	-2.38597	4.29E-16	KRT15	-3.82328	1.50E-06	PLAC9	-1.04045	4.10E-07

Figures

Figure 1

Normalize the three chips after quality control. A. The RLE plot of GSE7904; B. The signal distribution diagram of GSE7904; C. The RLE plot of GSE10797; D. The signal distribution diagram of GSE10797; E. The RLE plot of GSE10351; F. The signal distribution diagram of GSE10351

Figure 3

The DEGs heatmap of three datasets. A, B and C are GSE7904, GSE10797 and GSE103512. Each row represents an mRNA and each column represents a tumor sample. Clustering tree of the gene is shown on the left, and the sample clustering tree appears at the top. The color scale shown at the top illustrates the relative expression level of mRNA. Red color represents a high relative expression level and a green color represents a low relative expression level. D. The genetic alteration of the 7 genes.

Figure 4

Visualization of GO and KEGG. A. The Go-enrichment plot of GSE7904. B. The KEGG pathway analysis of DEGs from GSE7904. C. The Go-enrichment plot of GSE10797. D. The KEGG pathway analysis of DEGs from GSE10797. E. The Go-enrichment plot of GSE103512. F. The KEGG pathway analysis of DEGs from GSE103512

Figure 5

Survival analysis of the 7 DEGs

Figure 6

The expression level of 7 DEGs in Breast Cancer

Figure 7

Page 15/16

The drug targets gene of the 7 DEGs

Supplementary Files

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

- [FIGS1.jpg](#)
- [FIGS2.jpg](#)
- [FIGS3.jpg](#)