

Identification of candidate biomarkers correlated with poor prognosis of breast cancer based on bioinformatics analysis

Gang Chen^{a,*}, Mingwei Yu^{b,*}, Jianqiao Cao^a, Huishan Zhao^c, Yuanping Dai^d, Yizi Cong^a, and Guangdong Qiao^{a,*}

^aDepartment of Breast Surgery, The Affiliated Yantai Yuhuangding Hospital of Qingdao University, Yantai, Shandong, P.R. China;

^bDepartment of Orthopedics, The Affiliated Yantai Yuhuangding Hospital of Qingdao University, Yantai, Shandong, P.R. China; ^cReproductive Medicine Centre, The Affiliated Yantai Yuhuangding Hospital of Qingdao University, Yantai, Shandong, P.R. China; ^dDepartment of Medical Genetics, Liuzhou Maternal and Child Health Hospital, Guangxi, P.R. China

ABSTRACT

Breast cancer (BC) is a malignancy with high incidence among women in the world. This study aims to screen key genes and potential prognostic biomarkers for BC using bioinformatics analysis. Total 58 normal tissues and 203 cancer tissues were collected from three Gene Expression Omnibus (GEO) gene expression profiles, and then the differential expressed genes (DEGs) were identified. Subsequently, the Gene Ontology (GO) function and Kyoto Encyclopedia of Genes and Genome (KEGG) pathway were analyzed to investigate the biological function of DEGs. Additionally, hub genes were screened by constructing a protein–protein interaction (PPI) network. Then, we explored the prognostic value and molecular mechanism of these hub genes using Kaplan–Meier (KM) curve and Gene Set Enrichment Analysis (GSEA). As a result, 42 up-regulated and 82 down-regulated DEGs were screened out from GEO datasets. The DEGs were mainly related to cell cycles and cell proliferation by GO and KEGG pathway analysis. Furthermore, 12 hub genes (*FN1*, *AURKA*, *CCNB1*, *BUB1B*, *PRC1*, *TPX2*, *NUSAP1*, *TOP2A*, *KIF20A*, *KIF2C*, *RRM2*, *ASPM*) with a high degree were identified initially, among which, 11 hub genes were significantly correlated with the prognosis of BC patients based on the Kaplan–Meier-plotter. GSEA reviewed that these hub genes correlated with KEGG_CELL_CYCLE and HALLMARK_P53_PATHWAY. In conclusion, this study identified 11 key genes as BC potential prognosis biomarkers on the basis of integrated bioinformatics analysis. This finding will improve our knowledge of the BC progress and mechanisms.

ARTICLE HISTORY

Received 7 June 2021
Revised 21 July 2021
Accepted 22 July 2021

KEYWORDS

Bioinformatics; breast cancer; prognosis biomarker; gene expression omnibus; hub genes



Introductions

Breast cancer (BC) is the most common malignant tumor among women in the world and the leading cause to woman death [1], [2]. Alone United States up to 276,480 new women cases accounts for 30% of female cancers in the past years [3]. According to the molecular features of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (Her2), BC could be separated into four subtypes. Namely, Luminal A (ER+/PR+, Her2-); LuminalB (ER+/PR+, Her2+); HER2+ (ER-/PR-, Her2+); and triple-negative breast cancer (TNBC) (ER-/PR-, Her2-) [4]. Based on the different molecular types, different measures will be carried out during clinical treatment [5]. Despite modern advances in target therapy

method, the result of treating BC is still unsatisfactory because of drug resistance and recurrence. Thus, understanding the molecular mechanisms of BC and identifying novel potential prognostic biomarkers to improve the prognosis of BC are urgently needed.

Previous studies screened the biomarkers predictors and function enrichment, mostly using online tools [6,7], such as GEO2R, DAVID or KOBAS et al. Actually, the majority of the online database are not precise enough, due to the slow update. Besides, the online DE analysis tool GEO2R seems not to normalize original data, so there might be a great deviation for each probe.

In this study, we identified several key genes that could be used as sensitivity biomarkers for the diagnosis of BC based on Gene Expression

CONTACT Guangdong Qiao  qiaogdxy@163.com  Department of Breast Surgery, The Affiliated Yantai Yuhuangding Hospital of Qingdao University, NO. 20, Yudong Road, Yantai, Shandong, 264001, P.R. China

*These authors contributed equally to this work.

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Omnibus (GEO) database. We downloaded three different region source gene expression profiles (GSE29044, GSE42568, and GSE50428) from the GEO database consisting of 58 normal breast tissue samples and 203 BC tissue samples. Then, the limma package of R software and Venn diagram online tool were applied to differential expressed genes (DEGs) in the three datasets above. Furthermore, Gene Ontology (GO) function and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were conducted by the R software, and the newest annotation files were downloaded from the official website, respectively. We constructed a protein and protein interaction (PPI) network with Cytoscape and the hub genes were screened by the cytoHubba plugin. For the validation of the DEGs, the BC section in the database consisting of 1102 BC tissues and 113 normal tissues was downloaded using TCGA-Assembler [8] package of R software, and those DEGs from GEO were verified by the TCGA database. Finally, only 11 genes were selected as BC potential prognosis biomarkers by KM-plot, and GSEA analyses were involved to study the potential molecular mechanisms of these hub genes. To know our results reliability, we chose fewest reported non-hub genes and verified using BC patients sample, and it is consistent with our result. In conclusion, our study provides some potential sensitive biomarkers for BC patients and promotes an understanding of the molecular mechanisms of BC progression.

Materials and methods

Data source and processing

The GEO (<http://www.ncbi.nlm.nih.gov/geo/>) database is a free public database of microarrays and is used for gene expression datasets and platform records [9]. The gene expression profiles of GSE29044, GSE42568, GSE50428 were chosen from the GEO database. GSE29044 was based on the GPL570 platform, containing 36 normal breast tissues and 73 BC tissues. GSE42568 was based on the GPL570 platform, containing 17 normal breast tissues and 104 BC tissues. GSE50428 was based on the GPL13648 platform, containing 5 normal breast tissues and 26 BC tissues. The downloaded

data was studied using the Perl (Practical Extraction and Report Language, Version 5.30.2) software; then, log₂ transformation and Z score standardization were performed on all data of gene expression.

Identification of DEGs

DEGs between BC samples and normal breast samples were identified using the limma (version 3.30.0) package of R (version 3.5.1) software. The DEGs with $FC \geq 1.5$ or $FC \leq 1/1.5$ and adjust $P < 0.05$ were considered as the cutoff criteria. Then, we used Venn software online (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) to obtain the common DEGs in all three independent cohorts.

Functional enrichment analysis of DEGs

The GO annotation gives us a conspicuous meaningful for the variety of biological functional from microarray and other big datasets [10]. KEGG is a systematic gene and genomic function information database, which is stored in the PATHWAY segments [11]. GO and KEGG annotations were downloaded from the official website, respectively (<http://current.geneontology.org/products/pages/downloads.html>, https://www.genome.jp/kegg-bin/get_htext?hsa00001+3101, 2021-01-05 download), and enrichment of DEGs was performed using hypergeometric distribution formula of R software. We regarded $P\text{-value} \leq 0.05$ with fold change more than 2 as a statistically significant difference and significant enrichment.

PPI network construction and hub gene identification

After enrichment of the DEGs, we constructed the PPI network of all DEGs based on String online database (<https://string-db.org/>) [12] and Combined score greater than 0.9 as the cutoff criterion. The network was loaded to the Cytoscape (version 3.7.0) software, and cytoHubba plugin was carried out to predict hub genes [13].

Validation of the hub genes in the TCGA database

The TCGA_BRCA dataset was downloaded using the TCGA-Assembler (version 2.0.6) package of R software [8] and then the up- and down-regulated hub gene expression levels were verified by the DESeq package of R software. Then, the potential function of the hub genes was analyzed by the GSEA (version 4.1.0).

Breast cancer tissues

Seventeen breast cancer patients treated at the Department of Breast Surgery, The Affiliated Yantai Yuhuangding Hospital of Qingdao University were selected in our study. All the patients were signed informed consent forms approved by the Institutional Review Board of Yantai Yuhuangding Hospital.

RNA isolation and Q-PCR

Total RNA was isolated from BC patients tissue samples using Trizol reagent as the manufacturer's direction (Sparkjade, Qingdao, China). Following, 0.5 µg RNA from each sample was reversed to cDNA by the SPARK script II RT Plus Kit (Sparkjade, Qingdao, China). Then, Q-PCR was carried out using SYBR Green qPCR Mix kit (With ROX) (Sparkjade, Qingdao, China), following manufacturer's instructions. Finally, the expression level of mRNA was calculated by the $2^{-\Delta\Delta CT}$ formula.

Survival analysis

To further investigate the value of hub genes in breast cancer patients, the Kaplan–Meier plotter (<http://kmplot.com/analysis/>) analysis was conducted [7]. BC database was applied to estimate the prognosis values of hub genes. If the p-value ≤ 0.05 , it would be considered statistically significant.

Statistical analysis

All statistical analyses were performed using R software. The data was expressed as the

mean \pm standard deviation (SD) from the dataset. Statistical analyses were conducted two-tailed and none paired Student's t-test by the GraphPad Prism 5 (GraphPad Inc., San Diego, CA, USA) software, and $p \leq 0.05$ were considered as statistical significance [14].

Results

Identification of DEGs in BC

Three gene expression profiles (GSE29044, GSE42568, and GSE50428) were selected in this study. Data from each GEO data set were respectively analyzed using limma package of R software to screen DEGs ($FC \geq 1.5$ or $FC \leq 1/1.5$ and adjust $P < 0.05$). A total of 182, 705, 681 up-regulated and 351, 914, 944 down-regulated genes were filtered from GSE29044, GSE42568 and GSE50428, respectively. The volcano plot and heat map of DEGs were shown in (Figure 1(a-h)). In addition, 41 and 86 overlapped up and down-regulated DEGs were screened out by online Venn software from three gene expression profiles (Figure 1(d-e)).

GO functional and KEGG pathway enrichment analysis of DEGs

All DEGs enrichment score was calculated by the hypergeometric distribution of R software. For the up-regulated DEGs, Biological Process (BP) terms are most significantly enriched in cell division, negative regulation of B cell differentiation, and anaphase-promoting complex-dependent catabolic process. Cellular Component (CC) terms are most significantly enriched in spindle, collagen-containing extracellular matrix, and kinetochores. As for the Molecular Function (MF), extracellular matrix structural constituent, microtubule binding, and microtubule motor activity are mostly enriched (Figure 2(a-c)). Besides, for the down-regulated DEGs, BP terms are most significantly enriched in the cellular response to heparin, retinol metabolic process, and positive regulation of fat cell differentiation. CC terms are most significantly enriched in the extracellular space, collagen-containing extracellular matrix, and

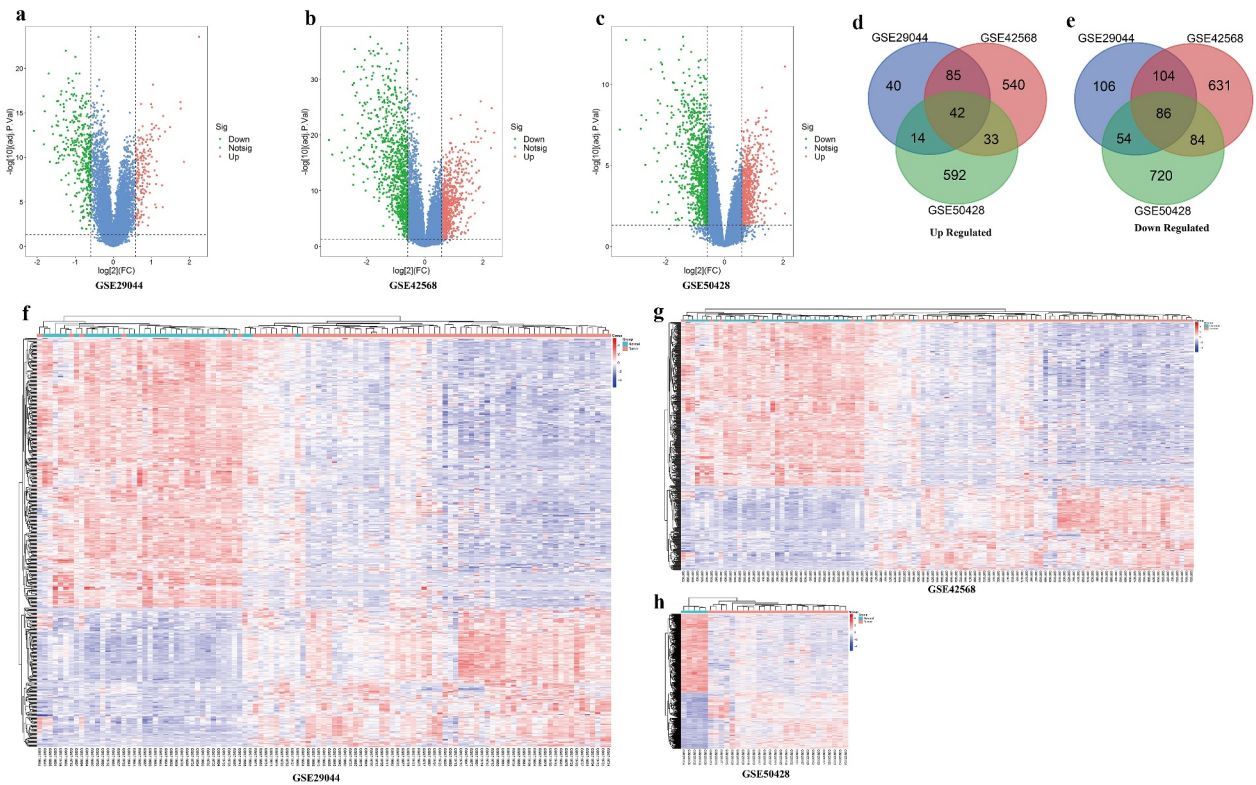


Figure 1. DEGs of three GEO profiles. (A-C, The volcano map of DEGs three GEO profiles. F-H, The heat map of DEGs three GEO profile. D, E, Venn diagram of three GEO profiles DEGs.)

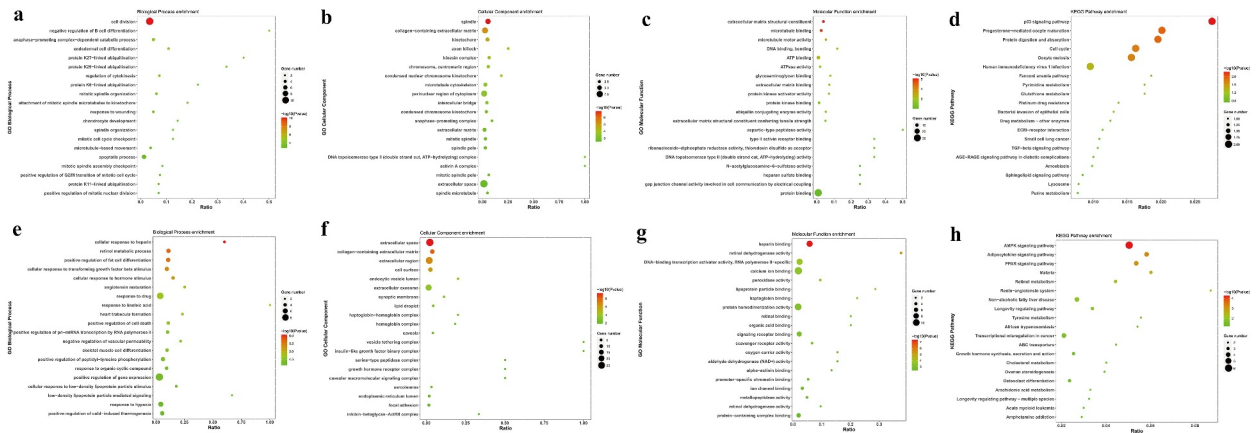


Figure 2. GO function and KEGG pathway analysis of DEGs. (A-D, GO function, and KEGG pathway analysis of up-regulated DEGs. E-H, GO function and KEGG pathway analysis of down-regulated DEGs.)

extracellular region. As for the MF, heparin binding, retinal dehydrogenase activity, and DNA-binding transcription activator activity, RNA polymerase II-specific are mostly enriched (Figure 2(e-g)). The method of enrichment is as same as that of the GO enrichment analysis. The up-regulated DEGs were most significantly

enriched in the p53 signaling pathway, progesterone-mediated oocyte maturation, protein digestion and absorption. As for the down-regulated DGEs, they are mainly enriched in the AMPK signaling pathway, Adipocytokine signaling pathway, and PPAR signaling pathway (Figure 2(d-h)).

PPI network construction and hub gene screening

The PPT network was constructed by the STRING online database and visualized by the Cytoscape software. The degree of the per node was calculated by the Cytohubba plugin, and the top 5 degrees of DEGs are considered as the hub genes of BC, 12 up-regulated (*FN1*, *AURKA*, *CCNB1*, *BUB1B*, *PRC1*, *TPX2*, *NUSAP1*, *TOP2A*, *KIF20A*, *KIF2C*, *RRM2*, *ASPM*) and 1 down-regulated (*PPARG*) (Figure 3).

DEGs validation

To further investigate the 13 candidate hub genes, we validated their expression in the TCGA_BRCA dataset. The expression levels of all these 13 hub genes were consistent with the results of GEO profiles analysis (Figure 4).

In the mean time, we chose least reported non-hub three genes (*CDC3A*, *ZWINT*, and *UBE2S*) and verified in BC samples by the Q-PCR. The expression levels of these three genes were consistent with GEO profiles (Figure 5).

Survival analysis of the identified hub genes

To deeply clarify the role of hub genes in BC patients prognosis, the overall survival of hub genes was analyzed using KM-plot (<http://kmplot.com/analysis/>). The KM plot showed that all the hub genes had a significant difference between high and low expression levels, except *FN1* (Figure 6). This result indicated that the 12 hub genes have prognostic significance for BC patients.

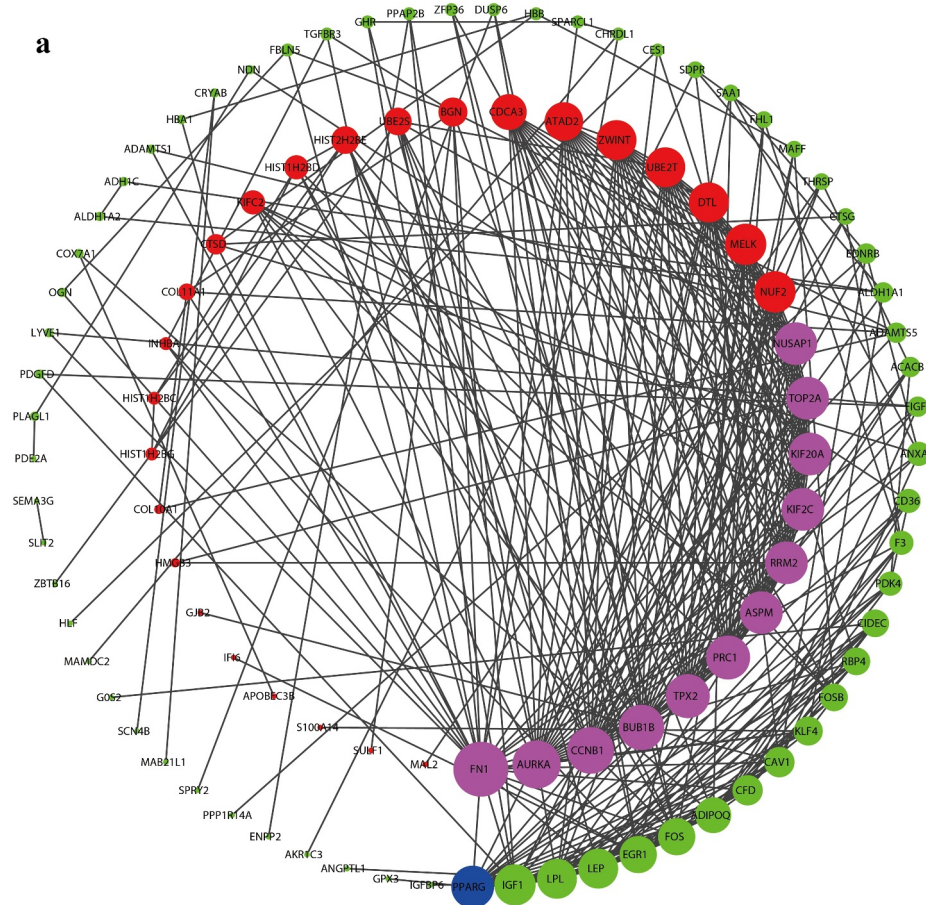


Figure 3. PPI analysis of hub genes of DEGs.

(A, Red color is up-regulated DEGs, purple is up-regulated hub genes, green is down-regulated genes, blue is down-regulated hub genes.)

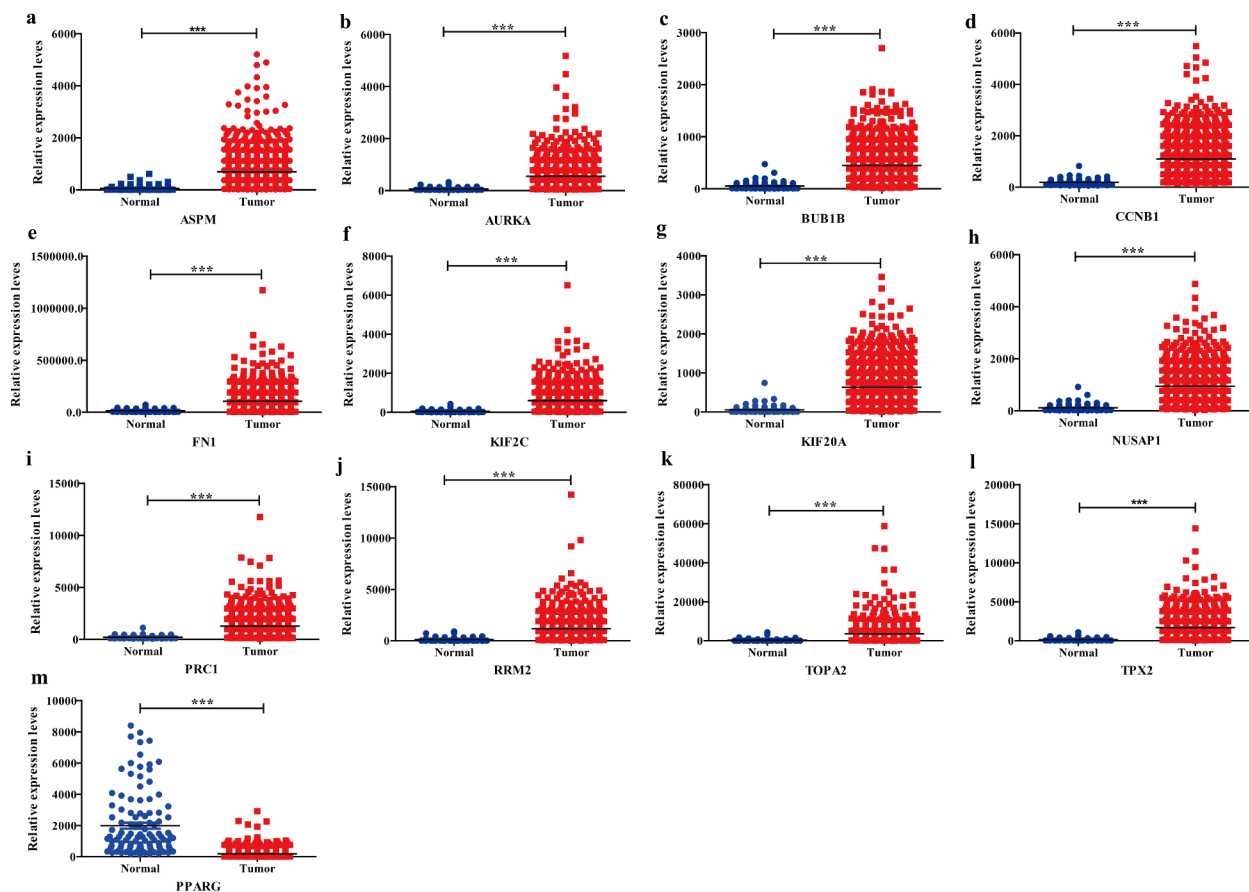


Figure 4. Expression validation of 13 hub targets in BC compared with adjacent tissues from TCGA data sets. (A-L, Up-regulated hub genes of BC from TCGA database. M, Down-regulated hub genes of BC from TCGA database.)

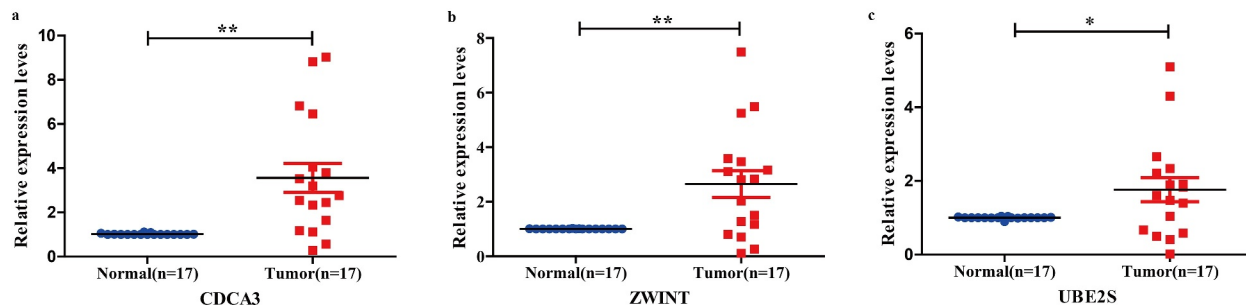


Figure 5. Non-hub genes validation in BC compared with adjacent tissues from patients. (A-C. CDC3A and ZWINT, UBE2S relative expression levels, respectively.)

Hub gene GSEA analysis

To evaluate the potential mechanism of these hub genes in BC, GSEA was performed to get the underlying pathway from the GSEA dataset. The expression levels of these genes from TCGA_BRCA were separated into two parts using median. From the GSEA results, we

could know that these hub genes were highly related to the ‘KEGG_CELL_CYCLE’ and ‘HALLMARK_P53_PATHWAY’ gene set (Figure 7). The result of GSEA indicated a significant difference ($FDR \leq 0.05$) in the enrichment of MSigDB Collection and was consistent with GO and KEGG enrichment.

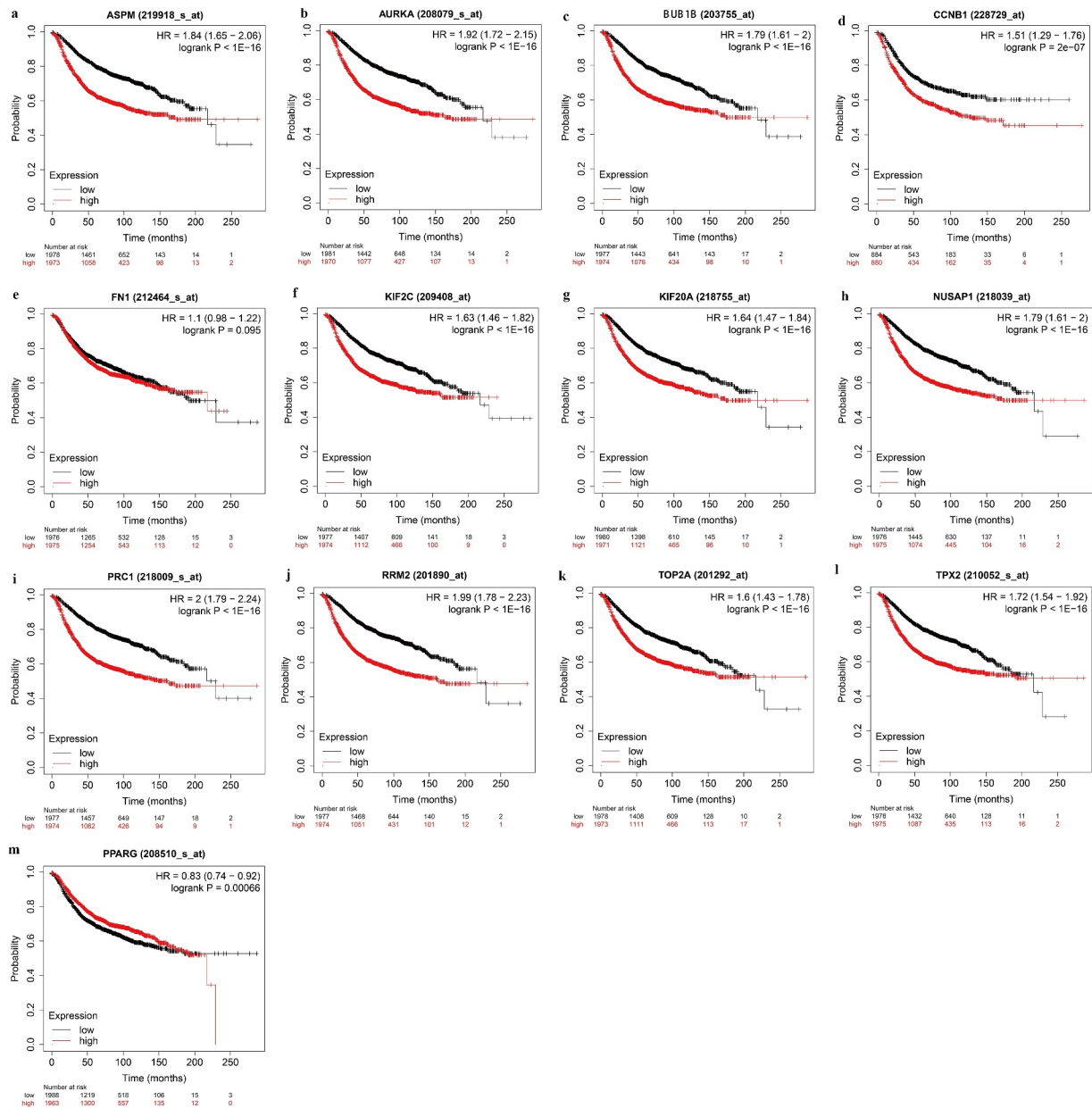


Figure 6. The prognostic gene signature of hub genes in the BC patients. (A-L, Overall survival of up-regulated hub genes. M, Overall survival of down-regulated hub genes.)

Discussion

As one of the most common cancers and the No.1 killer among women malignancies in the world, breast cancer brings women severe hurts physically and psychologically [15–17]. For the past ten years, the treatment of BC has many problems, such as drug resistance and poor prognosis, which reveals BC still is a high-risk disease. Therefore, exploring the molecular

pathogenesis of BC is crucial to fully understand the development and prognosis of BC is imminent. Although there are a lot of bioinformatics paper reported about BC, they are almost based on the online tool, DAVID database with slow update rate. Obviously, some of the results are incomprehensive. In our study, we download newest annotation from official website GO and KEGG. Tidied and calculated enrichment using

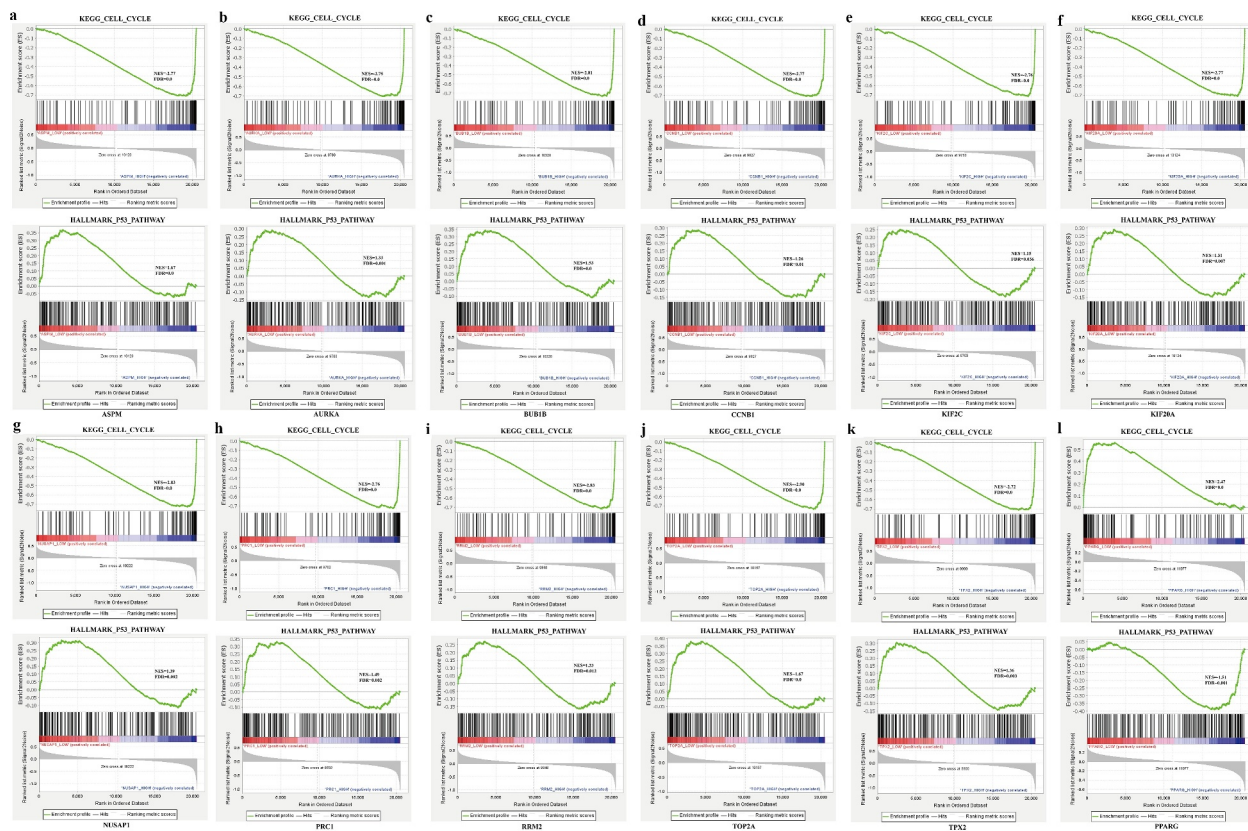


Figure 7. Enrichment plots from GSEA.

(A-K, Enrichment plots from GSEA of up-regulated hub genes. Enrichment plots from GSEA of down-regulated hub genes.)

hypertension formula. The results would be more considerate.

In this study, in order to select effective biomarkers of BC and identify their potential molecular mechanisms, we studied three GEO profiles and found 124 overlapping DEGs, containing 42 up-regulated genes and 86 down-regulated genes. After all the analyses we performed, 12 BC target genes (*AURKA*, *CCNB1*, *BUB1B*, *PRC1*, *TPX2*, *NUSAP1*, *TOP2A*, *KIF20A*, *KIF2C*, *RRM2*, *ASPM*, *PPARG*) were identified. The detailed information of these genes is as follows.

From the GO function enrichment, we could know that *AURKA*, *PRC1*, *BUB1B*, *CCNB1*, *ASPM*, and *KIF2C* hold the same function of cell division. *AURKA* is a serine/threonine kinase, which shares a highly conserved catalytic domain containing auto phosphorylating site [18]. It positively regulates cell cycle progression and plays a role in cell centrosome and spindle microtubules during mitosis [19]. On the other hand, it has been widely reported that *AURKA* is an oncogene to promote tumorigenesis in multiple types of cancer

including solid tumors (such as bladder cancer [20,21], prostate cancer [22,23], colon cancer [24]) and hematological malignancies [18]. *BUB1B* also belongs to serine/threonine-protein kinase and could lead to cell death and slow growth in BC cells [25]. The abnormal regulation of *PRC1* contributed to cancer progress [26,27], such as prostate cancer and breast cancer [28,29]. But, the mechanism of *PRC1* in cancer is still unclear. *CCNB1* is a regulatory protein involved in mitosis and a critical cell cycle regulator of the G2/M checkpoint [30]. Previous studies have reported that *CCNB1* could participate in oncogene pathways among many kinds of cancers, such as BC and colorectal cancer [31–34]. And at initial stage of cancer, it was more recognized by the T cells [7,35]. Abnormal spindle-like microcephaly associated gene (*ASPM*) encodes a protein of 3477 amino acids with an NH2-terminal microtubule-binding domain and two calponin homology domains [36]. On the one hand, *ASPM* is a regulator of Wnt and stemness in pancreatic adenocarcinoma [37] which as a Wnt associated

marker, it is not only could predict survival time but also could become a target therapy [38]. On the other hand, the function of *APSM* evidence has pointed out that it is an oncogene and its prognosis has been investigated in various cancers, such as epithelial ovarian cancer, gliomas, pancreatic and prostate cancer and liver cancer, as well as BC [37,39–43]. *MCAK* (also known as *KiF2C*), a member of the motor protein-13 motor family, is reported to undergo large conformational changes when it is converted from solution to microtubule-binding during its catalytic cycle [44]. The activity of *MCAK* is inhibited by Aurora B kinase through phosphorylation on multiple amino acids within its N-terminus [45,46]. Previous studies reported that the high expression of *KIF2C* could serve as an independent marker of poor prognosis in several tumors, including glioma, colorectal cancer, and gastric cancer [47–49], but the roles in BC reported less.

As for other hub genes, they were enriched in different GO BP terms. Targeting protein for *Xenopus* kinesin-like protein 2 (*TPX2*) is a microtubule-associated protein. The expression of *TPX2* is strictly regulated by the cell cycle. In general, *TPX2* exists in the G1 and S phases of the cell cycle and disappears at the end of cytokinesis [50]. A growing number of papers reported that the high expression of *TPX2* was connected with bad and shorter overall survival of patients in many tumors [51–53]. When using siRNA to knock down *TPX2*, the cycling-related proteins were down-regulated and cell apoptosis-related proteins were increased. It indicated that *TPX2* is an important cell signaling molecular [52]. However, the mechanism of *TPX2* in BC is still unknown. *NUSAP1* (encodes nucleolar and spindle-associated protein 1), a nucleolar spindle-associated protein, has been reported that plays a complicated and sensitized role in cell division and mitotic progression, spindle formation, and stability controlled by phosphorylation [54]. *NUSAP1* was highly expressed in kinds of malignancies and correlated with poor prognosis in aggressive triple-negative BC. *TOP2A* is a DNA topoisomerase that participates in many processes during transcription and replication through altering DNA topological structure [55]. A significantly high expression level of *TOP2A* has been reported

in many types of cancers [56–59]. And it was related to worse overall survival for various cancers [56,59]. Some researchers reported that *TOP2A* could induce apoptosis and suppress cell growth and invasion via AKT/ERK signaling pathways in colon cancer [7]. Kinesin family member 20A (*KIF20A*) is believed to modulate microtubule dynamics [60], which could promote the tumorigenesis and progression of prostate cancer and glioma [61], particularly the biochemical recurrence and metastasis [60,62]. Silencing *KIF20A* could induce prostate cancer cells to death and aberrantly activated *Gli2-KIF20A* axis which is crucial for the growth of hepatocellular carcinoma and predicts poor prognosis in hepatocellular carcinoma [63]. And *KIF20A* could induce paclitaxel resistance of BC [64]. Ribonucleotide reductase M2 subunit (*RRM2*), a rate-limiting enzyme involved in DNA synthesis and damage repair, plays important roles in many cellular processes such as cell growth, invasiveness, migration and senescence et al. [65]. *RRM2* as a tumor driver is frequently overexpressed in various malignancies [66–68]. Others found that the expression level of *RRM2* was correlated with invasion, cell differentiation, and metastasis in colorectal carcinoma [69], and correlated with lung cancer grade level [70]. Silencing *RRM2* attenuated melanoma growth, which was consistent with the maintenance of senescence-associated cell-cycle arrest [71]. Peroxisome proliferator-activated receptor gamma (*PPARG*) could induce cell cycle arrest, terminal differential, and anti-inflammatory [72,73], and induce G2/M cell cycle arrest by activating P38 in renal cancer and bladder cancer [74,75]. Furthermore, *PPARG* inducing the down-regulation of Wnt/beta-catenin pathway was observed and aberrant in many cancers [76]. Up-regulation of *PPARG* was correlated with downstream metabolic effectors. From a phenotypic point of view, this was associated with increased lung and lymph-nodes metastasis, indicating that the stratification method can be targeted for the treatment of aggressive diseases [77].

From the NCBI database, we could know that most of the hub genes were researched or verified, so we chose other non-hub genes and fewest reported three genes (*CDCA3*, *ZWINT*, *UBE2S*) to verified. These three genes were consistent

with predicted results. CDCA3 is a FBOX protein, and is essential for cell mitosis [78], which take part in forming E3 ligase complex and relate to ubiquitination. It triggers cells to enter mitosis and dephosphorylation of CDC2 proteins in G2 phase. Some reports showed that Skp1-Cullin-F-box complex controlled the G1/S phase and led to cancers if became imbalance [79,80]. It had been found that CDCA3 was significantly illustrated in the lung cancer samples and correlated with clinical progress [79]. And CDCA3 was associated with worse RFS and OS in Luminal A breast cancer [80]. Some evidence also suggested that CDCA3 interacting with TRAF2 could activate NF- κ B pathway in colorectal cancer. But, the mechanism of CDCA3 in BC is still absent. ZWINT has same function with CACA3, could interact with E3 ligase and promote cell growth [81]. Reported studies suggested that ZWINT was highly expressed in various cancers and related with cancer progress, such as lung cancer [82] and liver cancer [83]. From our result, it may have the same function in BC. UBE2S is a member of the E2 family of ubiquitin-conjugating enzymes and collaborates with anaphase-promoting complex to prolong K11-linkages and polyubiquitin chains on substrates for 26 S proteasome-mediated degradation [84–86]. Also, it is responsible for Lys11-linkage ubiquitin modifications on β -catenin [87]. Past years have reported that UBE2S is preternaturally expressed in kinds of cancers and negatively relates with patients progress [88–90]. Highly expressed UBE2S facilitates cell proliferation and metastasis via targeting suppressor of Von Hippel-Lindau and p53 degradation [90]. Knocking down UBE2S could inhibit EMT cell signaling and inhibits invasion of cervical cancer [91]. But, the mechanism of UBE2S in BC is still poorly understood.

In conclusion, our study identified 13 hub genes and 3 non-hub genes (*AURKA*, *CCNB1*, *BUB1B*, *PRC1*, *TPX2*, *NUSAP1*, *TOP2A*, *KIF20A*, *KIF2C*, *RRM2*, *ASPM*, *PPARG*, non-hub genes are *CDCA3*, *ZWINT* and *UBE2S*) that might be involved in the progression of BC using multiple gene expression data sets and a series of comprehensive analyses of bioinformatics. These findings provide new insights into the diagnosis and

treatment of the BC, while the main limitation of this research is lacking experiment to verify the hub genes expression level in the BC tissues and function in BC cells. Therefore, the further experimental studies are still needed to ensure our findings.

HIGHLIGHT

- (1) Identification of effective biomarkers of breast cancer.
- (2) *AURKA*, *CCNB1*, *BUB1B*, *PRC1*, *TPX2*, *NUSAP1*, *TOP2A*, *KIF20A*, *KIF2C*, *RRM2*, *ASPM*, *PPARG*, and *CDCA3*, *ZWINT*, *UBE2S* may could as a breast cancer target genes.
- (3) Cell division was mostly related.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Shandong Medical and Health Science and Technology Development Project [202004081034].

Author contributions

GC, MY, and JC conceived, designed and wrote the manuscript; YC and HZ analyzed the data; GQ and YD revised the manuscript; GC and MY contributed equally to this study; all authors read and approved the final manuscript.

References

- [1] Cheng L, Huang YZ, Chen WX, et al. Cell division cycle proteinising prognostic biomarker of breast cancer. *Biosci Rep*. 2020;40(5):BSR20191227.
- [2] Niu T, Zhang W, Xiao W. MicroRNA regulation of cancer stem cells in the pathogenesis of breast cancer. *Cancer Cell Int*. 2021;21(1):31.
- [3] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70:7–30.
- [4] Onitilo AA, Engel JM, Greenlee RT, et al. Breast cancer subtypes based on ER/PR and Her2 expression: comparison of clinicopathologic features and survival. *Clin Med Res*. 2009;7(1–2):4–13.
- [5] Yang G, Jian L, Chen Q. Comprehensive analysis of expression and prognostic value of the claudin family in human breast cancer. *Aging (Albany NY)*. 2021;13(6):8777–8796.
- [6] Lin Y, Fu F, Lv J, et al. Identification of potential key genes for HER-2 positive breast cancer based on

- bioinformatics analysis. *Medicine (Baltimore)*. 2020;99(1):e18445.
- [7] Liu S, Liu X, Wu J, et al. Identification of candidate biomarkers correlated with the pathogenesis and prognosis of breast cancer via integrated bioinformatics analysis. *Medicine (Baltimore)*. 2020;99(49):e23153.
- [8] Wei L, Jin Z, Yang S, et al. TCGA-assembler 2: software pipeline for retrieval and processing of TCGA/CPTAC data. *Bioinformatics*. 2018;34(9):1615–1617.
- [9] Zhao H, Jiang A, Yu M, et al. Identification of biomarkers correlated with diagnosis and prognosis of endometrial cancer using bioinformatics analysis. *J Cell Biochem*. 2020;121(12):4908–4921.
- [10] Gene Ontology Consortium. The Gene Ontology (GO) project in 2006. *Nucleic Acids Res*. 2006;34(90001):D322–6.
- [11] Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30.
- [12] Wang X, Yue J, Ren X, et al. Modularity analysis based on predicted protein-protein interactions provides new insights into pathogenicity and cellular process of *Escherichia coli* O157:H7. *Theor Biol Med Model*. 2011;8(1):47.
- [13] Song W, Ren J, Wang C, et al. Analysis of Circular RNA-related competing endogenous RNA identifies the immune-related risk signature for colorectal cancer. *Front Genet*. 2020;11:505.
- [14] Cai Y, Mei J, Xiao Z, et al. Identification of five hub genes as monitoring biomarkers for breast cancer metastasis in silico. *Hereditas*. 2019;156(1):20.
- [15] Baade P. Geographical variation in breast cancer outcomes. *Int J Environ Res Public Health*. 2017;14(5):523.
- [16] Liang Y, Zhang H, Song X, et al. Metastatic heterogeneity of breast cancer: molecular mechanism and potential therapeutic targets. *Semin Cancer Biol*. 2020;60:14–27.
- [17] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin*. 2019;69(1):7–34.
- [18] Du R, Huang C, Liu K, et al. Targeting AURKA in Cancer: molecular mechanisms and opportunities for Cancer therapy. *Mol Cancer*. 2021;20:15.
- [19] Carvalho S, Ribeiro SA, Arocena M, et al. The nucleoporin ALADIN regulates Aurora A localization to ensure robust mitotic spindle formation. *Mol Biol Cell*. 2015;26(19):3424–3438.
- [20] de Martino M, Shariat SF, Hofbauer SL, et al. Aurora A Kinase as a diagnostic urinary marker for urothelial bladder cancer. *World J Urol*. 2015;33(1):105–110.
- [21] Lei Y, Yan S, Ming-De L, et al. Prognostic significance of Aurora-A expression in human bladder cancer. *Acta Histochem*. 2011;113(5):514–518.
- [22] Matarasso N, Bar-Shira A, Rozovski U, et al. Functional analysis of the Aurora Kinase A Ile31 allelic variant in human prostate. *Neoplasia*. 2007;9(9):707–IN25.
- [23] Buschhorn HM, Klein RR, Chambers SM, et al. Aurora-A over-expression in high-grade PIN lesions and prostate cancer. *Prostate*. 2005;64(4):341–346.
- [24] Lentini L, Amato A, Schillaci T, et al. Aurora-A transcriptional silencing and vincristine treatment show a synergistic effect in human tumor cells. *Oncol Res*. 2008;17(3):115–125.
- [25] Koyuncu D, Sharma U, Goka ET, et al. Spindle assembly checkpoint gene BUB1B is essential in breast cancer cell survival. *Breast Cancer Res Treat*. 2021;185(2):331–341.
- [26] Brynychova V, Ehrlichova M, Hlavac V, et al. Genetic and functional analyses do not explain the association of high PRC1 expression with poor survival of breast carcinoma patients. *Biomed Pharmacother*. 2016;83:857–864.
- [27] Zhang B, Shi X, Xu G, et al. Elevated PRC1 in gastric carcinoma exerts oncogenic function and is targeted by piperlongumine in a p53-dependent manner. *J Cell Mol Med*. 2017;21(7):1329–1341.
- [28] Luo HW, Chen QB, Wan YP, et al. Protein regulator of cytokinesis 1 overexpression predicts biochemical recurrence in men with prostate cancer. *Biomed Pharmacother*. 2016;78:116–120.
- [29] Yun HJ, Cho YH, Moon Y, et al. Transcriptional targeting of gene expression in breast cancer by the promoters of protein regulator of cytokinesis 1 and ribonuclease reductase 2. *Exp Mol Med*. 2008;40(3):345–353.
- [30] Liu J, Wen Y, Liu Z, et al. VPS33B modulates c-Myc/p53/miR-192-3p to target CCNB1 suppressing the growth of non-small cell lung cancer. *Mol Ther Nucleic Acids*. 2021;23:324–335.
- [31] Fei F, Qu J, Liu K, et al. The subcellular location of cyclin B1 and CDC25 associated with the formation of polyploid giant cancer cells and their clinicopathological significance. *Lab Invest*. 2019;99(4):483–498.
- [32] Wang L, Xia Y, Chen T, et al. Sanyang Xuedai enhances the radiosensitivity of human non-small cell lung cancer cells via increasing iNOS/NO production. *Biomed Pharmacother*. 2018;102:618–625.
- [33] Monteiro LS, Diniz-Freitas M, Warnakulasuriya S, et al. Prognostic Significance of Cyclins A2, B1, D1, and E1 and CCND1 numerical aberrations in oral squamous cell carcinomas. *Anal Cell Pathol (Amst)*. 2018;2018:7253510.
- [34] Wang W, Shen XB, Jia W, et al. The p53/miR-193a/EGFR feedback loop function as a driving force for non-small cell lung carcinoma tumorigenesis. *Ther Adv Med Oncol*. 2019;11:1758835919850665.
- [35] Egloff AM, Weissfeld J, Land SR, et al. Evaluation of anticyclin B1 serum antibody as a diagnostic and prognostic biomarker for lung cancer. *Ann N Y Acad Sci*. 2005;1062(1):29–40.

- [36] Gao ZY, Yu F, Jia HX, et al. ASPM predicts poor prognosis and regulates cell proliferation in bladder cancer. *Kaohsiung J Med Sci.* 2020;36(12):1021–1029.
- [37] Wang WY, Hsu CC, Wang TY, et al. A gene expression signature of epithelial tubulogenesis and a role for ASPM in pancreatic tumor progression. *Gastroenterology.* 2013;145(5):1110–1120.
- [38] Hsu CC, Liao WY, Chan TS, et al. The differential distributions of ASPM isoforms and their roles in Wnt signaling, cell cycle progression, and pancreatic cancer prognosis. *J Pathol.* 2019;249(4):498–508.
- [39] Lin SY, Pan HW, Liu SH, et al. ASPM is a novel marker for vascular invasion, early recurrence, and poor prognosis of hepatocellular carcinoma. *Clin Cancer Res.* 2008;14(15):4814–4820.
- [40] Alsiary R, Bruning-Richardson A, Bond J, et al. Deregulation of microcephalin and ASPM expression are correlated with epithelial ovarian cancer progression. *PLoS One.* 2014;9(5):e97059.
- [41] Bikeye SN, Colin C, Marie Y, et al. ASPM-associated stem cell proliferation is involved in malignant progression of gliomas and constitutes an attractive therapeutic target. *Cancer Cell Int.* 2010;10(1):1.
- [42] Xie JJ, Zhuo YJ, Zheng Y, et al. High expression of ASPM correlates with tumor progression and predicts poor outcome in patients with prostate cancer. *Int Urol Nephrol.* 2017;49(5):817–823.
- [43] Turnbull AK, Arthur LM, Renshaw L, et al. Accurate prediction and validation of response to endocrine therapy in breast cancer. *J Clin Oncol.* 2015;33(20):2270–2278.
- [44] McHugh T, Zou J, Volkov VA, et al. The depolymerase activity of MCAK shows a graded response to Aurora B kinase phosphorylation through allosteric regulation. *J Cell Sci.* 2019;132(4):jcs228353.
- [45] Andrews PD, Ovechkina Y, Morrice N, et al. Aurora B regulates MCAK at the mitotic centromere. *Dev Cell.* 2004;6(2):253–268.
- [46] Lan W, Zhang X, Kline-Smith SL, et al. Aurora B phosphorylates centromeric MCAK and regulates its localization and microtubule depolymerization activity. *Curr Biol.* 2004;14(4):273–286.
- [47] Bie L, Zhao G, Wang YP, et al. Kinesin family member 2C (KIF2C/MCAK) is a novel marker for prognosis in human gliomas. *Clin Neurol Neurosurg.* 2012;114(4):356–360.
- [48] Ishikawa K, Kamohara Y, Tanaka F, et al. Mitotic centromere-associated kinesin is a novel marker for prognosis and lymph node metastasis in colorectal cancer. *Br J Cancer.* 2008;98(11):1824–1829.
- [49] Nakamura Y, Tanaka F, Haraguchi N, et al. Clinicopathological and biological significance of mitotic centromere-associated kinesin overexpression in human gastric cancer. *Br J Cancer.* 2007;97(4):543–549.
- [50] Heidebrecht HJ, Buck F, Steinmann J, et al. p100: a novel proliferation-associated nuclear protein specifically restricted to cell cycle phases S, G2, and M. *Blood.* 1997;90(1):226–233.
- [51] Hsu CW, Chen YC, Su HH, et al. Targeting TPX2 suppresses the tumorigenesis of hepatocellular carcinoma cells resulting in arrested mitotic phase progression and increased genomic instability. *J Cancer.* 2017;8(8):1378–1394.
- [52] Liang B, Jia C, Huang Y, et al. TPX2 level correlates with hepatocellular carcinoma cell proliferation, apoptosis, and EMT. *Dig Dis Sci.* 2015;60(8):2360–2372.
- [53] Liu HC, Zhang Y, Wang XL, et al. Upregulation of the TPX2 gene is associated with enhanced tumor malignancy of esophageal squamous cell carcinoma. *Biomed Pharmacother.* 2013;67(8):751–755.
- [54] Mills CA, Suzuki A, Arceci A, et al. Nucleolar and spindle-associated protein 1 (NUSAP1) interacts with a SUMO E3 ligase complex during chromosome segregation. *J Biol Chem.* 2017;292(42):17178–17189.
- [55] Tang Y, Zhang Y, Hu X. Identification of potential hub genes related to diagnosis and prognosis of hepatitis B virus-related hepatocellular carcinoma via integrated bioinformatics analysis. *Biomed Res Int.* 2020;2020:4251761.
- [56] Deng JL, Xu YH, Wang G. Identification of potential crucial genes and key pathways in breast cancer using bioinformatic analysis. *Front Genet.* 2019;10:695.
- [57] Zhang R, Xu J, Zhao J, et al. Proliferation and invasion of colon cancer cells are suppressed by knockdown of TOP2A. *J Cell Biochem.* 2018;119(9):7256–7263.
- [58] Liu HQ, Zhang SL, Song S. HER-2/neu and TOPIIa expression in gastric cancer reflect disease severity. *Hepatogastroenterology.* 2012;59:1290–1293.
- [59] Shen J, Yu S, Sun X, et al. Identification of key biomarkers associated with development and prognosis in patients with ovarian carcinoma: evidence from bioinformatic analysis. *J Ovarian Res.* 2019;12(1):110.
- [60] Yu H, Xu Z, Guo M, et al. FOXM1 modulates docetaxel resistance in prostate cancer by regulating KIF20A. *Cancer Cell Int.* 2020;20(1):545.
- [61] He Q, Meng J, Liu S, et al. Long non-coding RNA UCA1 upregulates KIF20A expression to promote cell proliferation and invasion via sponging miR-204 in cervical cancer. *Cell Cycle.* 2020;19(19):2486–2495.
- [62] Song Z, Huang Y, Zhao Y, et al. The identification of potential biomarkers and biological pathways in prostate cancer. *J Cancer.* 2019;10(6):1398–1408.
- [63] Shi C, Huang D, Lu N, et al. Aberrantly activated Gli2-KIF20A axis is crucial for growth of hepatocellular carcinoma and predicts poor prognosis. *Oncotarget.* 2016;7(18):26206–26219.
- [64] Khongkow P, Gomes AR, Gong C, et al. Paclitaxel targets FOXM1 to regulate KIF20A in mitotic catastrophe and breast cancer paclitaxel resistance. *Oncogene.* 2016;35(8):990–1002.
- [65] Nordlund P, Reichard P. Ribonucleotide reductases. *Annu Rev Biochem.* 2006;75(1):681–706.
- [66] Jin CY, Du L, Nuerlan AH, et al. High expression of RRM2 as an independent predictive factor of poor

- prognosis in patients with lung adenocarcinoma. Aging (Albany NY). 2020;13(3):3518–3535.
- [67] Grolmusz VK, Karasz K, Micsik T, et al. Cell cycle dependent RRM2 may serve as proliferation marker and pharmaceutical target in adrenocortical cancer. *Am J Cancer Res.* 2016;6:2041–2053.
- [68] Grossi F, Dal Bello MG, Salvi S, et al. Expression of ribonucleotide reductase subunit-2 and thymidylate synthase correlates with poor prognosis in patients with resected stages I-III non-small cell lung cancer. *Dis Markers.* 2015;2015:302649.
- [69] Lu AG, Feng H, Wang PX, et al. Emerging roles of the ribonucleotide reductase M2 in colorectal cancer and ultraviolet-induced DNA damage repair. *World J Gastroenterol.* 2012;18(34):4704–4713.
- [70] Hsu NY, Wu JY, Liu X, et al. Expression status of ribonucleotide reductase small subunits hRRM2/p53R2 as prognostic biomarkers in stage I and II non-small cell lung cancer. *Anticancer Res.* 2011;31:3475–3481.
- [71] Fatkhutdinov N, Sproesser K, Krepler C, et al. Targeting RRM2 and Mutant BRAF is a novel combinatorial strategy for melanoma. *Mol Cancer Res.* 2016;14(9):767–775.
- [72] Huang R, Zhang J, Li M, et al. The Role of Peroxisome Proliferator-Activated Receptors (PPARs) in Pan-Cancer. *PPAR Res.* 2020;2020:6527564.
- [73] Bandera Merchan B, Tinahones FJ, Macias-Gonzalez M. Commonalities in the Association between PPARG and Vitamin D Related with Obesity and Carcinogenesis. *PPAR Res.* 2016;2016:2308249.
- [74] Fujita M, Yagami T, Fujio M, et al. Cytotoxicity of troglitazone through PPARgamma-independent pathway and p38 MAPK pathway in renal cell carcinoma. *Cancer Lett.* 2011;312(2):219–227.
- [75] Plissonnier ML, Fauconnet S, Bittard H, et al. Insights on distinct pathways of thiazolidinediones (PPARgamma ligand)-promoted apoptosis in TRAIL-sensitive or -resistant malignant urothelial cells. *Int J Cancer.* 2010;127(8):1769–1784.
- [76] Vallee A, Lecarpentier Y, Vallee JN. Curcumin: a therapeutic strategy in cancers by inhibiting the canonical WNT/beta-catenin pathway. *J Exp Clin Cancer Res.* 2019;38(1):323.
- [77] Galbraith LCA, Mui E, Nixon C, et al. PPAR-gamma induced AKT3 expression increases levels of mitochondrial biogenesis driving prostate cancer. *Oncogene.* 2021;40(13):2355–2366.
- [78] Shen D, Fang Y, Zhou F, et al. The inhibitory effect of silencing CDCA3 on migration and proliferation in bladder urothelial carcinoma. *Cancer Cell Int.* 2021;21(1):257.
- [79] Adams MN, Burgess JT, He Y, et al. Expression of CDCA3 is a prognostic biomarker and potential therapeutic target in non-small cell lung cancer. *J Thorac Oncol.* 2017;12(7):1071–1084.
- [80] Perez-Pena J, Alcaraz-Sanabria A, Nieto-Jimenez C, et al. Mitotic read-out genes confer poor outcome in luminal A breast cancer tumors. *Oncotarget.* 2017;8(13):21733–21740.
- [81] Endo H, Ikeda K, Urano T, et al. Terf/TRIM17 stimulates degradation of kinetochore protein ZWINT and regulates cell proliferation. *J Biochem.* 2012;151(2):139–144.
- [82] Yuan W, Xie S, Wang M, et al. Bioinformatic analysis of prognostic value of ZW10 interacting protein in lung cancer. *Onco Targets Ther.* 2018;11:1683–1695.
- [83] Ying H, Xu Z, Chen M, et al. Overexpression of Zwint predicts poor prognosis and promotes the proliferation of hepatocellular carcinoma by regulating cell-cycle-related proteins. *Onco Targets Ther.* 2018;11:689–702.
- [84] Zhou G, Shen M, Zhang Z. ZW10 Binding Factor (ZWINT), a direct target of mir-204, predicts poor survival and promotes proliferation in breast cancer. *Med Sci Monit.* 2020;26:e921659.
- [85] Williamson A, Wickliffe KE, Mellone BG, et al. Identification of a physiological E2 module for the human anaphase-promoting complex. *Proc Natl Acad Sci U S A.* 2009;106(43):18213–18218.
- [86] Brown NG, Watson ER, Weissmann F, et al. Mechanism of polyubiquitination by human anaphase-promoting complex: RING repurposing for ubiquitin chain assembly. *Mol Cell.* 2014;56(2):246–260.
- [87] Li Z, Wang Y, Li Y, et al. Ube2s stabilizes beta-Catenin through K11-linked polyubiquitination to promote mesendoderm specification and colorectal cancer development. *Cell Death Dis.* 2018;9(5):456.
- [88] Yoshimura S, Kasamatsu A, Nakashima D, et al. UBE2S associated with OSCC proliferation by promotion of P21 degradation via the ubiquitin-proteasome system. *Biochem Biophys Res Commun.* 2017;485(4):483–498.
- [89] Ma Y, Li K, Li S, et al. Prognostic value of ubiquitin-conjugating enzyme E2 S overexpression in hepatocellular carcinoma. *Int J Biol Macromol.* 2018;119:225–231.
- [90] Pan YH, Yang M, Liu LP, et al. UBE2S enhances the ubiquitination of p53 and exerts oncogenic activities in hepatocellular carcinoma. *Biochem Biophys Res Commun.* 2018;503(2):895–902.
- [91] Lin TH, Hsu WH, Tsai PH, et al. Dietary flavonoids, luteolin and quercetin, inhibit invasion of cervical cancer by reduction of UBE2S through epithelial-mesenchymal transition signaling. *Food Funct.* 2017;8(4):1558–1568.