Bardet-Biedl Syndrome 4 in Early Diagnosis and Prognosis of Breast Cancer

H. ZHANG, B. SUO¹, XUE PU SUN, YING-XUAN SUN² AND H. Q. QIAO*

Department of General Surgery, The First Affiliated Hospital of Harbin Medical University, No. 25 Youzheng Street, Harbin, Heilongjiang 150008, ¹General Hospital of Heilongjiang Province Land Reclamation Bureau, Harbin 150088, ²Weifang Medical University, Weifang, Shandong 261053, China

Zhang et al.: Bardet-Biedl Syndrome 4 in Breast Cancer

Bardet-Biedl syndrome 4 is the key protein to control cilia formation. In this study, bioinformatics method was used to screen the core genes related to the prognosis of breast cancer by analyzing the gene chip data of gene expression omnibus and the cancer genome atlas database, so as to provide a new candidate target for the treatment of breast cancer. Data were downloaded from the cancer genome atlas, gene expression omnibus to evaluate Bardet-Biedl syndrome 4 expression levels in breast cancer. Differentially expressed genes were screened by R package. Gene ontology and Kyoto encyclopedia of Genes and Genomes pathway enrichment analysis was used to explore the biological functions of differentially expressed genes. The correlation of differentially expressed genes used was, “corrplot” for visual analysis. The protein-protein interaction relationship was constructed based on search tool for the retrieval of interacting genes/proteins database and Cytoscape software and the key genes were obtained by module analysis with Cytoscape software molecular complex detection plugin and the prognostic value and survival of Bardet-Biedl syndrome 4 were evaluated by R package. Finally, the correlation between Bardet-Biedl syndrome 4 and clinopathological parameters was also visualized by R package. These differentially expressed genes were mainly involved in response to peptide hormone, nuclear division, hormone secretion and transport and extracellular matrix. Genes were mainly involved in the Kyoto encyclopedia of genes and genomes pathway called Interleukin-17 signaling pathway. Bardet-Biedl syndrome 4 levels were found to be down regulated in breast cancer tissues compared with normal tissues. Survival analysis showed that low Bardet-Biedl syndrome 4 expression was associated with poor prognosis. These results were verified in clinical specimens, where in the Bardet-Biedl syndrome 4 protein levels were significantly down regulated in breast cancer tissues compared with non-breast cancer tissues. This study confirmed that Bardet-Biedl syndrome 4 can be used as an independent prognostic factor for the prognosis of breast cancer, which provides a basis for exploring a new target for the treatment of breast cancer.

Key words: Bardet-Biedl syndrome 4, breast cancer, bioinformatics, differentially expressed genes, prognosis

Breast cancer (BC) is one of the malignant tumors with the highest incidence in women and it is also the main cause of cancer death among women in the world[1]. Globally, with the development of screening and treatment, the survival rate of patients with BC has been improved. However, due to the high genetic heterogeneity of BC, some types of BC cells also have a high degree of invasion and metastasis and are easy to relapse[2]. Therefore, traditional treatments such as radiotherapy and surgery can improve patients but cannot completely prevent patients from dying of cancer metastasis. In order to develop better treatment strategies and better understand the targets of its potential mechanism, we urgently need to explore more specific and economical biomarkers to predict the prognosis of BC.

Bardet-Biedl syndrome (BBS) is a rare multisystem disease characterized by retinal dystrophy, renal insufficiency, obesity, polydactyly and many developmental and behavioral defects. It is usually inherited by autosomal recessive inheritance, which is caused by homozygous or compound heterozygous mutations[3]. So far, 24 genes associated with BBS have been identified (BBS1-BBS24). Eight of them formed

*Address for correspondence
E-mail: qiaohaiquan@hotmail.com
a stable protein complex, called BBSome, which is a binding protein complex composed of BBS1, 2, 4, 5, 7, 8, 9 and 18, which is structurally similar to the coat protein or adapter protein involved in vesicle transport[4-6]. Bardet-Biedl syndrome 4 (BBS4) is one of several proteins that cause BBSome. BBS4 is involved in the regulation of transmembrane proteins, including the transport of G protein-coupled receptor (GPCR) into the cilia and the output of extra ciliated receptors, regulation of cell cycle progression and microtubule tissue during cell division[7]. Recent studies have shown that silent BBS4 induces accelerated cell division and abnormal differentiation. Therefore, in this study, BBS4 was selected as the research object to determine its exact role and mechanism in tumor cell migration and invasion.

The purpose of this study was to evaluate the prognostic and therapeutic value of BBS4 as a potential biomarker of BC. We compare its performance with other biomarkers and gene signatures currently available. Based on the analysis of the downloaded Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) data sets, we found that BBS4 has prognostic value in patients with BC.

MATERIALS AND METHODS

Data sources

The microarray dataset GSE42568 was downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). GSE42568 contained gene expression profile data of 104 cases of BC and 17 cases of normal breast biopsies and was performed on GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array by Clarke C et al. The analysis data of BBS4 in BC progression and prognosis was downloaded from TCGA (https://portal.gdc.cancer.gov/). All data sets needed to be normalized using the R package before analysis.

Fundamental analysis

Firstly, the annotation file of the corresponding platform was used to annotate the gene expression profile and then use the “limma” package of Bioconductor (http://www.bioconductor.org/packages/release/bioc/html/limma.html) to normalize the expression data[8]. Secondly, a ranked list was generated according to the fold changes of the up-regulated and down-regulated genes in each data set. Finally, use the “robust rank aggregation” R package to perform integrated analysis on these sorted gene lists[9]. p values of each gene were calculated and the final gene list was sorted in descending order of p value. In addition, Bonferroni correction was used to reduce false positive results.

Function and pathway enrichment analysis

Using the R software to estimate differentially expressed genes (DEGs) between BC and non-BC samples, we performed Gene ontology (GO) term enrichment analysis, including molecular function (MF), cellular component (CC) and biological process (BP), as well as the Kyoto Encyclopedia of Genes and Genomes (KEGG). DEGs were identified using “clusterProfiler” (http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html), “org.Hs.eg.db” (http://www.bioconductor.org/packages/release/data/annotation/html/org.Hs.eg.db.html), “enrichplot” (http://www.bioconductor.org/packages/release/bioc/html/enrichplot.html) and “ggplot2” packages in Bioconductor, which is an R package that researchers can use to perform GO functional enrichment analysis and KEGG pathway enrichment analysis. Probe sets with multiple gene symbols or without corresponding gene symbols were deleted. Genes with two or more probe sets were absolute maximized. The cut-off criteria were set as follows: p<0.05 and logFC (fold change)>1.

Protein-protein interaction (PPI) network and module analysis

Protein-protein interaction (PPI) network of DEGs in BC can be constructed from the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) 11.0 database (http://string-db.org/cgi/input.pl) with a confidence level of 0.04. Then, Cytoscape 3.7.2 software was used to visualize the PPI network and the Molecular Complex Detection (MCODE) plugin of Cytoscape software was used to check module analysis to screen key genes[10].

Statistical analysis

All statistical analysis data were processed using R software (R version 4.0.2) unless otherwise specified. Student t-test and one-way analysis of variance were applied to analysis of continuous data and Pearson’s Chi-square test was used in the processing of categorical data. Kaplan-Meier survival analysis was used to plot overall survival (OS). The hazard ratio (HR) of the 95% confidence intervals (CI) will be computed using Cox proportional hazard regression analysis to test the relationship between BBS4 expression level and
RESULTS AND DISCUSSION

In order to identify the DEGs related to BC and non-BC, differential expression analysis was effectuated using the R software package and a list of up-regulated and down-regulated DEGs was obtained. Using logFC (fold change)>1 and p value<0.05 as the standard, 145 DEGs were screened out, of which 91 were up regulated genes and 54 were down regulated genes (Table 1). These DEGs hierarchical clustering graph and volcano map are shown in fig. 1.

The GO function enrichment analysis and KEGG pathway enrichment analysis were performed on 145 DEGs through R software. GO analysis includes BP, CC and MF. The results of GO analysis showed that the main BP of the selected DEGs involved in organelle fission, regulation of chromosome separation, response to peptide hormone, nuclear division, hormone secretion and transport, positive regulation of secretion. The primary CC terms consisted of extracellular matrix, condensed chromosome outer kinetochore, anaphase promoting complex, nuclear chromosome, centromeric region. Finally, the primary MF terms included epidermal growth factor receptor binding (Table 2 and fig. 2A). The KEGG analysis consisted of the DEGs were mainly concentrated in cell cycle and Interleukin-17 (IL-17) signaling pathway (Table 2 and fig. 2B). For instance, BBS4 may participate in multicellular organismal homeostasis (GO:0048871), regulation of blood pressure (GO:0008217), positive regulation of developmental growth (GO:0048639), cellular response to leptin stimulus (GO:0044320), response to leptin (GO:0044321), regulation of multicellular organism growth (GO:0040014), epithelial tube morphogenesis (GO:0060562), heart morphogenesis (GO:0003007), retina homeostasis.

TABLE 1: DEGs (91 UP-REGULATED GENES AND 54 DOWN-REGULATED GENES)

<table>
<thead>
<tr>
<th>DEGs</th>
<th>Genes name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up-regulated</td>
<td>BBS4, GFRA1, CTSF, NME5, TBC1D9, GSTM2, SEC14L2, MAPT, ADIRF, GSTM3, GAJ1, AKR7A3, ARSG, STC2, IGF1R, DNAL1, XBP1, NPY1R, RNAS34, CA12, IGFBP4, TTC39A, QDPR, TGFR3, ELOVL2, CPE, ACADSB, ABAT, AG2R, ZBTB20, CFD, PAK2, SLC39A6, GPA, FOXA1, ERBB4, MYB, SHC2, KAL1, THBS4, DACH1, PLA2G16, LRR1C17, COX7A1, GHR, ADC4Y1, IRS1, SEPP1, DCN, AGTR1, SERPIN1, EFEMP1, FMD, SPARCL1, GRIA2, PSD3, ARNT2, HSPA2, RTN1, PPP1R3C, TFF1, TFF3, ADIPOQ, PCAT18, FBPI, GRP, LEP, A2GPI, CLGN, CRIP1, AREG, DUSP4, SCGB2A2, ADH1B, SERPIN5, PLIN1, SCN1A, FABP4, SERPINA3, KRT19, PLAT, KCNE4, TNNT1, CYP2B7P, MYBPC1, UGT2B28, DHR52, SCGB1D2, CPB1, HMGCS2, PIP, TTK, FOXM1, MELK, SLC7A5, MYBL2, BUB1, UBE2C, NDC80, EZH2, SLC25A43, CNE1, CDC25B, PTTG1, S100A8, KMO, BUB1B, RARRES51, TPX2, CDC20, PRKX, DGBP5, RM2, APOBEC3B, S100A9, DSC2, SOX11, AIM2, LAMP3, BCL2A1, TRIP13, ANP32E, KCNN4, LCN2, PLA2G7, MPP1, MPP9, MARCO, FGB, BBOX1, CDH3, CXCL10, CT83, ASS1, PROM1, ZIC1, MX1, MMP12, SLPI, S100A7, IFI44L, CXCL9, CHI3L1, CRISP3, GABRP</td>
</tr>
<tr>
<td>Down-regulated</td>
<td>TTK, FOXM1, MELK, SLC7A5, MYBL2, BUB1, UBE2C, NDC80, EZH2, SLC25A43, CNE1, CDC25B, PTTG1, S100A8, KMO, BUB1B, RARRES51, TPX2, CDC20, PRKX, DGBP5, RM2, APOBEC3B, S100A9, DSC2, SOX11, AIM2, LAMP3, BCL2A1, TRIP13, ANP32E, KCNN4, LCN2, PLA2G7, MPP1, MPP9, MARCO, FGB, BBOX1, CDH3, CXCL10, CT83, ASS1, PROM1, ZIC1, MX1, MMP12, SLPI, S100A7, IFI44L, CXCL9, CHI3L1, CRISP3, GABRP</td>
</tr>
</tbody>
</table>

Fig. 1: Heat map and volcanic maps of 145 DEGs (including 91 up-regulated genes and 54 down-regulated genes) from GSE42568 (A) Rows represent the genes, columns represent the samples. Red: up-regulation and Blue: down-regulation. (B) Black plots: normally expressed messenger RNAs; red plots: up-regulated DEGs; green plots: down-regulated DEGs.
### TABLE 2: GO AND KEGG PATHWAY ENRICHMENT ANALYSIS OF GENES IN THREE SIGNIFICANT MODULES

<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>Count</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOTERM_BP</td>
<td>GO:0048285—organelle fission</td>
<td>16</td>
<td>9.65E-09</td>
</tr>
<tr>
<td>GOTERM_BP</td>
<td>GO:0043434—response to peptide hormone</td>
<td>15</td>
<td>4.63E-08</td>
</tr>
<tr>
<td>GOTERM_BP</td>
<td>GO:0000280—nuclear division</td>
<td>14</td>
<td>1.35E-07</td>
</tr>
<tr>
<td>GOTERM_BP</td>
<td>GO:0051047—positive regulation of secretion</td>
<td>13</td>
<td>1.58E-06</td>
</tr>
<tr>
<td>GOTERM_BP</td>
<td>GO:1901987—regulation of cell cycle phase transition</td>
<td>13</td>
<td>5.53E-06</td>
</tr>
<tr>
<td>GOTERM_BP</td>
<td>GO:0046879—hormone secretion</td>
<td>12</td>
<td>3.01E-07</td>
</tr>
<tr>
<td>GOTERM_BP</td>
<td>GO:0009914—hormone transport</td>
<td>12</td>
<td>4.10E-07</td>
</tr>
<tr>
<td>GOTERM_BP</td>
<td>GO:0051306—mitotic sister chromatid separation</td>
<td>8</td>
<td>2.39E-09</td>
</tr>
<tr>
<td>GOTERM_CC</td>
<td>GO:0062023—extracellular matrix</td>
<td>12</td>
<td>3.75E-06</td>
</tr>
<tr>
<td>GOTERM_CC</td>
<td>GO:0000940—condensed chromosome outer kinetochore</td>
<td>12</td>
<td>6.37E-05</td>
</tr>
<tr>
<td>GOTERM_CC</td>
<td>GO:0005680—anaphase-promoting complex</td>
<td>3</td>
<td>2.26E-04</td>
</tr>
<tr>
<td>GOTERM_CC</td>
<td>GO:000780—condensed nuclear chromosome, centromeric regional activity</td>
<td>3</td>
<td>4.33E-04</td>
</tr>
<tr>
<td>GOTERM_MF</td>
<td>GO:0005154—epidermal growth factor receptor binding</td>
<td>4</td>
<td>4.86E-05</td>
</tr>
<tr>
<td>KEGG_PATHWAY</td>
<td>hsa04110:Cell cycle</td>
<td>7</td>
<td>1.42E-04</td>
</tr>
<tr>
<td>KEGG_PATHWAY</td>
<td>hsa04657:IL-17 signaling pathway</td>
<td>6</td>
<td>2.24E-04</td>
</tr>
</tbody>
</table>

GO: Gene Ontology; BP: biological process; CC: cellular component; MF: molecular function; KEGG: Kyoto Encyclopedia of Genes and Genomes. The count was the number of enriched genes in each term. If there were >5 terms enriched in this category, the top five terms were selected according to their p value.

![Fig. 2: GO analysis and KEGG pathway analysis of DEGs identified by R software](image)

(A) The GO terms of the BP, CC and MF categories enrichment of the 145 DEGs. BP: biological process; CC: cellular component; MF: molecular function. (B) The KEGG pathway analysis of the 145 DEGs.
Maximum Depth=100). A total of 6 central nodes were identified out of 78 nodes, then the central node of BBS4 was shown in fig. 4B and they were all up-regulated genes.

The expression level of BBS4 gene in the 1085 BC group and 291 normal control group in the TCGA database was analyzed online through the web page Gene Expression Profiling Interactive Analysis (GEPIA) and the results showed that the expression level of BBS4 in the BC group was lower than the normal control group (p<0.01) (fig. 5A). Then we evaluated the grade score for each patient according to the grade assessment model. In order to determine whether the grade score as an independent prognostic indicator, we performed univariate and multivariate analyses and evaluated the prognostic value of grade scores. The results of univariate analysis revealed that grade score was significantly associated with OS time (fig. 5B). Multivariate analysis showing the grade score as an independent prognostic factor in BC (fig. 5C). We evaluated the survival of the high grade group and grade risk group by the Kaplan-Meier survival curve, results showed that patients in the high-risk group has significantly longer survival time than those in the low-risk group (fig. 5D).

The correlation between BBS4 expression and the age, tumor size and tumor grade of BC patients was analyzed by R software (fig. 6). As the tumor grade was higher, the expression level of BBS4 was also higher (p<0.05), but there was no correlation with age and tumor size (p>0.05).

In order to further evaluate the predictive accuracy of BBS4, receiver operating characteristic (ROC) analysis was used to evaluate the sensitivity and specificity of BBS4 markers in predicting prognosis. As shown in
Fig. 5: Construction of nomogram
(A) The expression level of BBS4 gene in BC group was lower than that in normal control group. (B) The results of univariate analysis revealed that grade score was significantly associated with OS time. (C) Multivariate analysis showing the grade score as an independent prognostic factor in BC. (D) The survival curve of BBS4

Fig. 6: Correlation diagram between BBS4 and case parameters
(A) There was no correlation between BBS4 expression and the age of BC patients (p>0.05). (B) There was no correlation between BBS4 expression and the size of BC patients. (C) As the grade was higher, the expression level of BBS4 was also higher (p<0.05)
fig. 7, the area under the curve is 0.832, indicating that BBS4 has relatively high sensitivity and specificity. Therefore, along with the above results, BBS4 can be used as a potential marker to predict the prognosis and survival of patients with BC and has important clinical value.

BC is a heterogeneous disease, which can be divided into four molecular subtypes according to the expression of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) and proliferation marker Ki67 (MKI67). In recent years, the incidence of BC has been increased year by year and the age of onset was getting younger. Early diagnosis and treatment can improve the prognosis of patients. Therefore, the search for sensitive biomarkers is of great significance for the treatment and prognosis of patients with BC. This study combines the sequencing data of microarray chips for the first time to evaluate the potential value of BBS4 in BC. Single chip analysis showed that the low expression of BBS4 in BC was related to poor prognosis (p<0.05). Signal pathway analysis of co-expressed genes found that BBS4 may affect some cancer signal pathways, especially IL-17 signal pathway, leading to the occurrence and development of BC.

There are few studies on the expression of BBS4 in various tumor tissues. According to previous studies, it has been confirmed that the increased expression of p62/sequestosome 1 (SQSTM1) promotes the aggregation and dislocation of BBS4, which is a key protein that controls cilia formation. P62/SQSTM1 has been identified as a nuclear factor erythroid 2-related factor 2 (Nrf2) target gene, which is a key autophagy binding protein. Tetsuya et al. suggested that phosphorylated p62-dependent Nrf2 activation inhibitors can suppress the proliferation of hepatocellular carcinoma (HCC) and the resistance to anticancer agents. This Nrf2 inhibitor can be used to reduce the resistance of cancer cells to anticancer drugs. Ichimura et al. founded that p62-Kelch-like ECH-associated protein 1 (Keap1)-Nrf2 pathway can induce tumorigenesis of precancerous cells and promote tumor cell growth and drug resistance through metabolic reprogramming mediated by Nrf2 activation. In addition, many reports have shown that p62-Keap1-Nrf2 pathway plays a protective role under oxidation and stress conditions. For example, apoptosis induced by endoplasmic reticulum stress can be prevented by activating p62-Keap1-Nrf2 pathway; Quercetin attenuates hepatotoxicity induced by various hepatotoxic agents by activating p62-Keap1-Nrf2 pathway. Di et al. demonstrated that IL-17 signaling pathway also perpetuates BC metastasis.

In view of the fact that the expression and biological function of BBS4 in BC are not clear, this study first analyzed the differential genes between BC and normal breast tissues by bioinformatics using gene chip GSE42568 of GEO database. 145 DEGs were identified, including 91 up-regulated genes and 54 down-regulated genes. Subsequently, GO annotation and KEGG pathway analysis were performed to elucidate the significance of these identified aberrantly expressed genes. This enrichment analysis showed that the DEGs were mainly enriched in the pathways of cell movement and IL-17 signaling pathway. It is reported in the literature that the effect of IL-17 on cancer has been widely studied and it has different effects on different cancers. IL-17, was an important proinflammatory cytokine. Chen et al. confirmed that neutrophils, Gamma delta (γδ) T cells and IL173 have metastasis promoting effects in BC patients. Seth demonstrated that IL-17 signaling pathway also perpetuates BC metastasis. IL-17 was thought to be a key event in tumor progression and was associated with tumor metastasis. Subsequently, the data of TCGA were used to analyze the expression of BBS4 in patients with BC and its value in diagnosis and prognosis. The results confirmed that the expression of BBS4 was significantly decreased in BC. ROC curve showed that the expression level of BBS4 was of moderate diagnostic value in BC and Kaplan-Meier analysis showed that the expression level of BBS4 was correlated with grade. In addition, multivariate analysis

![Fig. 7: ROC curve analysis reveals the sensitivity and specificity of BBS4 in predicting the OS of patients](image-url)
confirmed that the expression of BBS4 can be used as an independent risk factor for the prognosis of BC.

Taken together, this study showed that BBS4 gene was down-regulated in BC tissues compared with normal tissues and its expression level was related to the prognosis of BC patients. Therefore, as a human tumor suppressor gene, the expression level of BBS4 may be a reliable auxiliary parameter for diagnosing BC and predicting tumor metastasis and prognosis.

Conflicts of interest

The authors report no conflicts of interest.

REFERENCES


This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

This article was originally published in a special issue, "Evolutionary Strategies in Biomedical Research and Pharmaceutical Sciences" Indian J Pharm Sci 2021:83(3) Spl issue;145-152