Nomogram Models Based on Gene Expression in Prediction of Breast Cancer Bone Metastasis

Teng-di Fan

Ningbo Medical Center Lihuili Hospital

Di-kai Bei

Ningbo Medical Center Lihuili Hospital

Song-wei Li (✉ lisongweitg@163.com)

Ningbo Medical Center Lihuili Hospital

Research Article

Keywords: Breast cancer, bone metastasis, gene signature, prognosis

Posted Date: November 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1109324/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Objective: To design a weighted co-expression network and build gene expression signature-based nomogram (GESBN) models for predicting the likelihood of bone metastasis in breast cancer (BC) patients.

Methods: Dataset GSE124647 was used as a training set, and GSE14020 was taken as a validation set. In the training cohort, limma package in R was adopted to obtain differentially expressed genes (DEGs) between BC non-bone metastasis and bone metastasis patients, which were used for functional enrichment analysis. After weighted co-expression network analysis (WGCNA), univariate Cox regression and Kaplan-Meier plotter analyses were performed to screen potential prognosis-related genes. Then, GESBN models were constructed and evaluated. Further, the expression levels of genes in the models were explored in the training set, which was validated in GSE14020. Finally, the prognostic value of hub genes in BC was explored.

Results: A total of 1858 DEGs were obtained. WGCNA result showed that the blue module was most significantly related to bone metastasis and prognosis. After survival analyses, GJA1, SLC24A3, ITGBL1, and SLC44A1 were subjected to construct a GESBN model for overall survival. While GJA1, IGFBP6, MDFI, ITGFBI, ANXA2, and SLC24A3 were subjected to build a GESBN model for progression-free survival. Kaplan-Meier plotter and receiver operating characteristic analyses presented the reliable prediction ability of the models. Besides, GJA1, IGFBP6, ITGBL1, SLC44A1, and TGFBI expressions were significantly different between the two groups in GSE124647 and GSE14020. The hub genes had a significant impact on patient prognosis.

Conclusion: Both the four-gene signature and six-gene signature could accurately predict patient prognosis, which may provide novel treatment insights for BC bone metastasis.

Introduction

Breast cancer (BC) is one of the most prevalent malignancies and the major cause of cancer-associated deaths of women worldwide [1]. BC is considered to have the highest diagnostic rate in cancer, with more than 1.6 million new cases detected a year, accounting for approximately one-quarter of all cancers in women [2]. Despite the substantial improvements in prognosis have been achieved due to better therapeutic approaches over the past 20 years, BC is still one of the most common causes of cancer deaths among females. According to statistics, a total of 522,000 women died from BC in 2012, while 627,000 individuals died from this disease in 2018 [3, 4].

Nearly 70% of BC patients developed bone metastasis, leading to osteolytic and osteoblastic cancers [5]. Tumor cells secreted factors including parathyroid hormone in the bone to create an environment conducive to osteolysis instead of the direct destruction of bone [6]. In addition, bone metastasis often contributes to adverse events such as hypercalcemia, spinal cord compression, fractures, and pain, which severely affect the quality of life in BC patients [7]. Magnetic resonance imaging, computed tomography,

X-ray are conventional imaging methods to detect bone metastasis but fail to sense tiny tumor masses and negligible tumor-induced osteolysis [8]. Besides, various metastatic bone lesions which resulted from BC are hard to eradicate by adjuvant localized radiotherapy or surgical intervention [9]. Therefore, clinicians are faced with a big challenge in the diagnosis and treatment of bone metastatic BC.

Currently, pathologic diagnosis has been considered as the gold standard for diagnosis and pathology-based classification, which is critical in guiding the treatment of BC patients [10]. Nevertheless, gene expression also represents an essential role in the prognosis of patients, thus providing clinically relevant information and targeted therapies [11, 12]. Hence, it's imperative to find novel biomarkers to improve the early diagnosis and prognosis of patients with bone metastatic BC by using public databases of gene expression data.

In this study, 1858 differentially expressed genes (DEGs) between BC non-bone metastasis and BC bone metastasis were obtained, of which 76 genes were probably related to bone metastasis and prognosis based on weighted co-expression network analysis (WGCNA). With univariate Cox regression and Kaplan-Meier plotter analyses, a four-gene expression signature-based nomogram model for overall survival (OS) and a six-gene expression signature-based nomogram model for progression-free survival (PFS) were constructed, respectively. Kaplan-Meier plotter and receiver operating characteristic (ROC) analyses validated the potential value of the models for predicting patient survival. Finally, the expression of the genes in nomogram models and the prognostic value of hub genes were initially explored and validated.

Materials And Methods

Data mining from gene expression omnibus (GEO) database

GEO database (<https://www.ncbi.nlm.nih.gov/geo/>) was used to obtain the BC microarray dataset by setting the following filter: (1) more than 50 samples with bone metastasis information; (2) with survival data; (3) with expression profiling data. Finally, GSE124647 dataset was chosen as a training set to identify the DEGs between non-bone metastasis and bone metastasis samples. The platform was Affymetrix Human Genome U133A Array (GPL96). In total, there were 140 samples containing clinical and RNA-seq expression data in the GSE124647. Besides, 65 samples in GSE14020 dataset were used as validation cohorts to verify the expression levels of key genes. Normalized gene expression was measured as log2-based transformation.

Identification and functional enrichment analysis of DEGs

The R package limma was used to screen the DEGs between BC non-bone metastasis and BC bone metastasis groups. $|\log_2 \text{FC}| > 1$ and P-value < 0.05 were set as the filtering parameters. Then, gene ontology (GO) including biological process (BP), cellular component (CC), and molecular function (MF), and KEGG were carried out to determine the major biological functions of these DEGs in the database for annotation, visualization, and integrated discovery (DAVID) (<https://david.ncifcrf.gov/summary.jsp>). P < 0.05 was considered statistically significant.

WGCNA

WGCNA is a systemic method that uses gene expression data to build a scale-free network [13]. A weighted co-expression network with the expression profile data of the DEGs was built using the WGCNA package of R. Following this, we screened the key module related to BC bone metastasis and prognosis, and then extracted the genes for further analysis.

Nomogram model construction and model effectiveness evaluation

Using the “survival” package in R, univariate Cox regression analysis was performed to obtain the potential prognostic genes based on OS and PFS. Only genes that had a significant impact on OS or PFS were considered to pass univariate Cox regression analysis screening. In addition, the prognostic value of the significant genes obtained in the univariate Cox regression analysis was evaluated by Kaplan-Meier plotter analysis. Only genes with statistical significance in OS or PFS analyses were considered to pass the screening. The intersected genes generated in univariate Cox regression and Kaplan-Meier plotter analyses were then entered into the construction of nomogram models in terms of OS and PFS using the “rms” package in R. The calibration curves were drawn to measure the performance of the two models. The genes which had the greatest contribution were selected as hub genes.

After that, the patients were divided into high-risk or low-risk groups using the optimal cut-off value of risk score, which was calculated by the “maxstat” package in R. And Kaplan-Meier plotter analyses were adopted to assess the survival difference between the two groups using “survfit” function of “survival” package in R. Moreover, the time-independent ROC analyses were conducted to further evaluate the prognostic value of the nomogram models by using the “pROC” package in R. $P < 0.05$ was considered as significantly different. The area under curve (AUC) was used as an indicator of prognostic accuracy.

The expression levels of prognostic genes in nomogram models

The expression levels of key genes between BC non-bone metastasis and BC bone metastasis groups in GSE124647 were firstly explored using t-test. Then, GSE14020 as a validation dataset was used to assess the differential expression of the key genes in two groups.

Validation of the prognostic value of the hub genes

Kaplan-Meier plotter (<http://kmplot.com/analysis/index.php?p=background>) is capable to assess the effect of 54,000 genes on survival in 21 cancer types. We used this database to verify the prognostic significance of the hub genes in BC. Survival curves were generated by the Kaplan-Meier method using the log-rank test. Log-rank P value less than 0.05 was statistically significant.

Results

Identification and functional enrichment of DEGs

Taking BC non-bone metastasis samples as a control group, 1858 DEGs in the training set including 992 up-regulated and 866 down-regulated genes were generated according to the selection criteria. The volcano plot and heat map of the DEGs were presented in **Figure 1A**, and **Figure 1B**, respectively.

To have a biological understanding of these DEGs, they were subjected to the DAVID database for GO annotation and KEGG pathway enrichment analysis. The top enriched GO terms in BPs were signal transduction, positive regulation of transcription from RNA polymerase II promoter, and immune response, and those in CCs were cytoplasm, cytosol, and extracellular exosome (**Figure 2A-2B**). The major MFs were protein binding, Poly(A) RNA binding, and identical protein binding (**Figure 2C**). In the KEGG pathway enrichment analysis, these genes were mainly involved in the MAPK signaling pathway, proteoglycans in cancer, and focal adhesion (**Figure 2D**).

WGCNA

We incorporated the expression profile of integrated DEGs with clinical traits of the BC samples to construct a gene co-expression network. Clinical characteristics including sample group, PFS time, OS time, OS status, and PFS status were clustered with expression matrix (**Figure 3A**). Then, we chose the optimal $\beta=6$ to ensure that network was scale-free (β was a soft-thresholding parameter that could emphasize strong correlations between genes and penalize weak correlations). After choosing the power of 2, the adjacency was transformed into a topological overlap matrix (TOM), which could measure the network connectivity of a gene defined as the sum of its adjacency with all other genes for the network gene ration, and the corresponding dissimilarity (1-TOM) was calculated (**Figure 3B**). Based on TOM, the average linkage hierarchical clustering was conducted to cluster genes by setting the minimum number of genes for each gene network module to 30. To further analyze the module, we calculated the eigengenes of each module and merged the modules by setting a height of 0.25. Finally, a total of 4 modules were acquired (**Figure 3C-3D**). The genes in the grey module could not be incorporated into any other module. Next, Pearson's correlation coefficients of the module eigengene of each module and the sample characteristics were calculated. The blue module with 76 genes was closely related to bone metastasis and survival status (**Figure 3E**). Thus, the genes in the blue module were chosen for further analysis.

Construction of the gene expression signature-based nomogram (GESBN) model

Univariate Cox regression and Kaplan-Meier plotter analyses were carried out on 140 patients in the GSE124647 to evaluate the association of 76 gene expression profiles in the blue module with patient OS and PFS. In univariate Cox regression analysis, significant genes related to OS were SLC44A1, SLC24A3, PDGFC, ITGBL1, and GJA1 (**Figure 4A**) (all $P < 0.05$). Eleven genes including MDFI, IGFBP6, GJA1, ANXA2, SLC24A3, TGFB1, CELA2A, CELA2B, CLEC11A, PPEF2, and SLC44A1 were notably linked to PFS (**Figure**

4B) (all $P < 0.05$). However, only four genes related to OS, and six genes related to PFS with statistical differences were extracted in Kaplan-Meier plotter analysis (**Figure 5 and Figure 6**) (all $P < 0.05$). Taken together, GJA1, SLC24A3, ITGBL1, and SLC44A1 were defined as potential prognostic genes for OS. GJA1, IGFBP6, MDFI, ITGFBI, ANXA2, and SLC24A3 were potential genes correlated with PFS. These prognostic genes were then subjected to the construction of nomogram models based on OS and PFS (**Figure 7A-7B**). For OS, SLC44A1 had the greatest contribution, which could reach 100 points, while MDFI contributed most to PFS. Therefore, SLC44A1 and MDFI were considered as hub genes. To ensure the accuracy of the GESBN models, the calibration curves were drawn. The calibration curves showed good agreement between prediction and observation in the probability of 1-, 3- and 5- year survival (**Figure 7C-7D**), indicating that the accuracy of the nomogram model was reliable.

Evaluation of the GESBN models

The GESBN score was calculated for each patient in the training set. Patients were ranked based on their risk scores and assigned into two groups as high-risk and low-risk of bone metastases. Using the survival package in R for survival analysis, the results showed that the OS rate of patients in the high-risk group was low, and the difference between the two groups was statistically significant (**Figure 8A**) ($P < 0.001$). Similarly, an unfavorable PFS was observed in the high-risk group patients (**Figure 8B**) ($P < 0.001$), suggesting that two nomogram models could predict survival well. Further, the time-independent ROC curves were drawn using the pROC package in R. In terms of OS, the AUCs of the 1-, 3-, and 5-year survival rates were 0.68, 0.62, and 0.72, respectively (**Figure 8C**). For PFS, the AUCs of the 1-, 3-, and 5-year survival rates were 0.73, 0.88, 0.94, respectively (**Figure 8D**), indicating that nomograms had a good predictive ability.

The expression levels of genes in nomogram models

Due to the predictive ability of nomogram models for both OS and PFS, we explored the expression levels of these key genes. In the training dataset of GSE124647, the expression levels of all the prognosis-related genes were significantly different between control and bone-metastasis groups (**Figure 9A**) (all $P < 0.05$). In the validation dataset of GSE14020, GJA1, IGFBP6, ITGBL1, SLC44A1, and ITGFBI expressions in the bone-metastasis group were different from those in the control group. The differences were statistically significant (**Figure 9B**) ($P < 0.05$).

Validation of the prognostic value of hub genes

Based on the nomogram result, SLC44A1 and MDFI were the hub genes. Kaplan-Meier plotter was performed to verify the effect of SLC44A1 and MDFI on OS, PFS, and DMFS in BC. Patients in the high SLC44A1 expression group tended to have favorable OS, PFS, and DMFS (**Figure 10A-10C**) ($P < 0.05$). Although MDFI expression was not significantly linked to OS and DMFS of the BC patients (**Figure 10D**,

10F) ($P > 0.05$), its high expression predicted worse PFS (**Figure 10E**) ($P < 0.01$). These results indicated that SLC44A1 and MDFI might be potential biomarkers for BC.

Discussion

BC is a heterogeneous tumor driven by various molecular progression pathways [14]. Analyses of BC progression showed that bone is the first metastatic site of this disease possibly due to the favorable chemokine milieu or microenvironment in the bone, as well as the intrinsic molecular characteristics of cancer cells [15, 16]. Although these hypotheses are promising, biological information and anatomical characteristics are still the basis for clinicians to determine prognosis; however, the predictors of bone metastasis remain uncertain clinically. [17, 18]. Some gene signature-based prognostic prediction models for BC patients have been reported via repurposing and analysis of microarray data [19, 20]. These models were built for predicting OS for BC patients but lack the prediction of bone metastasis. By using GEO accession number GSE124647, we obtained 1858 DEGs between BC non-bone metastasis and bone metastasis groups. After screening the prognosis-related genes, we constructed a four-gene expression signature-based nomogram model and a six-gene expression signature-based nomogram model.

We firstly conducted a differential analysis of the GSE124647 dataset in relation to BC bone metastasis and employed functional enrichment analysis to these DEGs, which were found to be mainly related to signal transduction, and positive regulation of transcription in terms of BP. CCs were mainly enriched in cytoplasm, and cytosol. MFs were mainly protein binding, and Poly(A) RNA binding. The potential pathways that they were involved in were MAPK signaling pathway and proteoglycans in cancer. Based on WGCNA, 76 genes in the blue module were initially selected for the following prognostic analysis. After univariate Cox regression and Kaplan-Meier plotter analyses, OS nomogram including GJA1, SLC24A3, ITGBL1, and SLC44A1, and PFS nomogram including GJA1, IGFBP6, MDFI, TGFBI, ANXA2, and SLC24A3 were constructed. Then, its reliable prognostic ability for the OS and PFS was confirmed by Kaplan-Meier plotter and ROC analyses. AUC can be used to assess the accuracy and predictive capacity of biomarkers in diagnostic tests [21].

After confirming the predictive value of nomograms, we validated the prognostic significance of hub genes in the models. We found that high SLC44A1 expression was significantly related to favorable OS, PFS, and DMFS. However, patients in the high MDFI group predicted worse PFS. The solute carrier (SLC) superfamily contains various membrane-bound transporters which are required to transport a wide variety of substrates over biological membranes, and the dysregulated expression of these transporters may be related to cancer metastasis [22]. SLC01B1 was found to be highly expressed in colon cancer, and its expression level was significantly associated with the degree of differentiation in this type of cancer [23]. SLC01B3 overexpression may be linked to hormone-dependent growth mechanisms, and the expression of this transporter could serve as a valid prognostic factor for BC [24]. As a member of SLC superfamily, SLC44A1 is a mitochondrial protein mediating choline transport, and is preferentially expressed in neurons and oligodendrocytes [25]. Besides, high activity of the SLC44A1 promoter has been proved to participate in the occurrence of papillary glioneuronal tumors [26]. Our study revealed an

important finding that high SLC4A1 expression contributed to the favorable clinical outcome for BC metastasis patients. MDFI is a transcription factor that negatively regulates myogenic family proteins [27]. Previous study has demonstrated that the loss of MDFI was related to human BC and myeloid neoplasm via negative regulation of Wnt pathway [28]. In this study, high MDFI expression led to poor PFS for patients with metastatic BC.

Due to the limited clinical characteristics of patients included in the GSE124647

GSE14020, we failed to verify the independent prognostic value of nomograms. Besides, experiments were not conducted due to various objective reasons, which would be implemented when the conditions are available in the future. Despite of these limitations, our study employed the bioinformatic approaches to assess the potential genes related to BC bone metastasis and predict the prognosis of metastatic BC patients.

Conclusion

Based on the construction of a weighted co-expression network for DEGs between BC non-bone metastasis and bone metastasis, we screened the key module and related genes to investigate a prognostic nomogram model for bone metastatic BC. The study provided some potent biomarkers of BC bone metastasis and enable the prediction of patient survival. We also found that SLC4A1 and MDFI were the hub genes in BC bone metastasis, which might be the therapeutic targets for this disease.

Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: The dataset used and/or analyzed during the current study is available from the corresponding author on reasonable request.

Competing interests: The authors have no conflicts of interest to declare.

Funding: No funding was received for conducting this study.

Author contributions: TD is mainly responsible for conception and design, DK is mainly responsible for data collection and assembly, and SW is mainly responsible for data analysis and interpretation. All authors wrote manuscript together. All authors read and approved the final manuscript.

Acknowledgement: Not applicable.

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, et al. 2015**Global cancer statistics, 2012.***CA Cancer J Clin*, **65**:87-108. doi:10.3322/caac.21262.
2. Jemal A, Bray F, Center MM, Ferlay J, Ward E, et al. 2011**Global cancer statistics.***CA Cancer J Clin*, **61**:69-90. doi:10.3322/caac.20107.
3. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, et al. 2015**Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012.***Int J Cancer*, **136**:E359-386. doi:10.1002/ijc.29210.
4. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, et al. 2018**Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.***CA Cancer J Clin*, **68**:394-424. doi:10.3322/caac.21492.
5. Coleman RE. 2006**Clinical features of metastatic bone disease and risk of skeletal morbidity.***Clin Cancer Res*, **12**:6243s-6249s. doi:10.1158/1078-0432.CCR-06-0931.
6. Chirgwin JM, Guise TA. 2000**Molecular mechanisms of tumor-bone interactions in osteolytic metastases.***Crit Rev Eukaryot Gene Expr*, **10**:159-178.
7. Gralow J. 2005**Evolving role of bisphosphonates in women undergoing treatment for localized and advanced breast cancer.***Clin Breast Cancer*, **5 Suppl**:S54-62. doi:10.3816/cbc.2005.s.005.
8. Xing H, Bu W, Zhang S, Zheng X, Li M, et al. 2012**Multifunctional nanoprobe for upconversion fluorescence, MR and CT trimodal imaging.***Biomaterials*, **33**:1079-1089. doi:10.1016/j.biomaterials.2011.10.039.
9. Albain KS, Swann RS, Rusch VW, Turrisi AT, 3rd, Shepherd FA, et al. 2009**Radiotherapy plus chemotherapy with or without surgical resection for stage III non-small-cell lung cancer: a phase III randomised controlled trial.***Lancet*, **374**:379-386. doi:10.1016/S0140-6736(09)60737-6.
10. Prat A, Pineda E, Adamo B, Galvan P, Fernandez A, et al. 2015**Clinical implications of the intrinsic molecular subtypes of breast cancer.***Breast*, **24 Suppl 2**:S26-35. doi:10.1016/j.breast.2015.07.008.
11. Li B, Cui Y, Diehn M, Li R. 2017**Development and Validation of an Individualized Immune Prognostic Signature in Early-Stage Nonsquamous Non-Small Cell Lung Cancer.***JAMA Oncol*, **3**:1529-1537. doi:10.1001/jamaoncol.2017.1609.
12. Bing Z, Tian J, Zhang J, Li X, Wang X, et al. 2016**An Integrative Model of miRNA and mRNA Expression Signature for Patients of Breast Invasive Carcinoma with Radiotherapy Prognosis.***Cancer Biother Radiopharm*, **31**:253-260. doi:10.1089/cbr.2016.2059.
13. Liu S, Song A, Wu Y, Yao S, Wang M, et al. 2021**Analysis of genomics and immune infiltration patterns of epithelial-mesenchymal transition related to metastatic breast cancer to bone.***Transl Oncol*, **14**:100993. doi:10.1016/j.tranon.2020.100993.
14. Esteva FJ, Sahin AA, Cristofanilli M, Arun B, Hortobagyi GN. 2002**Molecular prognostic factors for breast cancer metastasis and survival.***Semin Radiat Oncol*, **12**:319-328. doi:10.1053/srao.2002.35251.
15. Kang Y, Siegel PM, Shu W, Drobnjak M, Kakonen SM, et al. 2003**A multigenic program mediating breast cancer metastasis to bone.***Cancer Cell*, **3**:537-549. doi:10.1016/s1535-6108(03)00132-6.

16. Jones DH, Nakashima T, Sanchez OH, Kozieradzki I, Komarova SV, et al. 2006**Regulation of cancer cell migration and bone metastasis by RANKL.***Nature*, **440**:692-696. doi:10.1038/nature04524.
17. Hess KR, Pusztai L, Buzdar AU, Hortobagyi GN. 2003**Estrogen receptors and distinct patterns of breast cancer relapse.***Breast Cancer Res Treat*, **78**:105-118. doi:10.1023/a:1022166517963.
18. Millar EK, Graham PH, O'Toole SA, McNeil CM, Browne L, et al. 2009**Prediction of local recurrence, distant metastases, and death after breast-conserving therapy in early-stage invasive breast cancer using a five-biomarker panel.***J Clin Oncol*, **27**:4701-4708. doi:10.1200/JCO.2008.21.7075.
19. Li J, Wang W, Xia P, Wan L, Zhang L, et al. 2018**Identification of a five-lncRNA signature for predicting the risk of tumor recurrence in patients with breast cancer.***Int J Cancer*, **143**:2150-2160. doi:10.1002/ijc.31573.
20. Iuliano A, Occhipinti A, Angelini C, De Feis I, Lio P. 2018**Combining Pathway Identification and Breast Cancer Survival Prediction via Screening-Network Methods.***Front Genet*, **9**:206. doi:10.3389/fgene.2018.00206.
21. Hanley JA, McNeil BJ. 1982**The meaning and use of the area under a receiver operating characteristic (ROC) curve.***Radiology*, **143**:29-36. doi:10.1148/radiology.143.1.7063747.
22. Sutherland R, Meeson A, Lowes S. 2020**Solute transporters and malignancy: establishing the role of uptake transporters in breast cancer and breast cancer metastasis.***Cancer Metastasis Rev*, **39**:919-932. doi:10.1007/s10555-020-09879-6.
23. Pressler H, Sissung TM, Venzon D, Price DK, Figg WD. 2011**Expression of OATP family members in hormone-related cancers: potential markers of progression.***PLoS One*, **6**:e20372. doi:10.1371/journal.pone.0020372.
24. Muto M, Onogawa T, Suzuki T, Ishida T, Rikiyama T, et al. 2007**Human liver-specific organic anion transporter-2 is a potent prognostic factor for human breast carcinoma.***Cancer Sci*, **98**:1570-1576. doi:10.1111/j.1349-7006.2007.00570.x.
25. Michel V, Bakovic M. 2012**The ubiquitous choline transporter SLC44A1.***Cent Nerv Syst Agents Med Chem*, **12**:70-81. doi:10.2174/187152412800792733.
26. Bridge JA, Liu XQ, Sumegi J, Nelson M, Reyes C, et al. 2013**Identification of a novel, recurrent SLC44A1-PRKCA fusion in papillary glioneuronal tumor.***Brain Pathol*, **23**:121-128. doi:10.1111/j.1750-3639.2012.00612.x.
27. Elabd C, Ichim TE, Miller K, Anneling A, Grinstein V, et al. 2018**Comparing atmospheric and hypoxic cultured mesenchymal stem cell transcriptome: implication for stem cell therapies targeting intervertebral discs.***J Transl Med*, **16**:222. doi:10.1186/s12967-018-1601-9.
28. Cigognini D, Corneo G, Fermo E, Zanella A, Tripputi P. 2007**HIC gene, a candidate suppressor gene within a minimal region of loss at 7q31.1 in myeloid neoplasms.***Leuk Res*, **31**:477-482. doi:10.1016/j.leukres.2006.09.007.

Figures

Figure 1

Differentially expressed genes between breast cancer non-bone metastasis and bone metastasis patients from the GSE124647 dataset. (A). The volcano plot of the 992 up-regulated (red triangle) genes and 866 down-regulated (green triangle) genes. (B) The heat map of the top 50 significant DEGs.

Figure 2

Functional enrichment analysis of differentially expressed genes. (A) Biological process. (B) Cellular component. (C) Molecular function. (D) KEGG pathway.

Figure 3

Weighted co-expression network analysis. (A) Dendrogram of sample clustering and heatmap of clinical traits of all breast cancer samples in a dataset of GSE124647. (B) The optimal β value result graph. (C) Module eigengene dendrogram. The horizontal axis represents a color block, and each of the different color blocks represents a different module, and the vertical axis represents the height of the dendrogram based on the expression value. (D) Clustering of module eigengenes. (E) The correlation between gene modules and sample characteristics.

Figure 4

Forest plot of univariate Cox regression analyses of prognosis. (A). Significant genes related to overall survival. (B). Significant genes related to progression-free survival.

Figure 5

Prognostic genes related to overall survival by Kaplan-Meier plotter. Abbreviations: L, Low; H, High; HR, Hazard ratio.

Figure 6

Prognostic genes related to progression-free survival by Kaplan-Meier plotter. Abbreviations: L, Low; H, High; HR, Hazard ratio.

Figure 7

Construction and calibration of gene expression signature-based model. (A) The four-gene-based nomogram model on overall survival. (B) The six-gene-based nomogram model on progression-free survival. (C). Calibration curve for predicting patient overall survival at 1-, 3- and 5-year. (D). Calibration curve for predicting patient progression-free survival at 1-year.

Figure 8

Prognostic evaluation of nomogram models in GSE124647. Kaplan-Meier plotter (A). overall survival curve and (B). progression-free survival curve between high-risk and low-risk bone metastatic patients. (C). Receiver operating characteristic (ROC) analysis of the four-gene signature. (D). ROC analysis of the six-gene signature. Abbreviations: L, Low; H, High; HR, Hazard ratio.

Figure 9

Expression levels of genes in nomograms. (A). GSE 124647. (B). GSE14020.

Figure 10

Prognostic value of the hub genes in breast cancer. The effect of SLC44A1 on (A). overall survival, (B). progression-free survival, and (C). distant metastasis-free survival in breast cancer patients. The effect of MDFI on (D). overall survival, (E). progression-free survival, and (F) distant metastasis-free survival. Abbreviations: HR, Hazard ratio.