Systematic Review and Meta-Analysis

Medicine

OPEN

Screening and identification of hub-gene associated with brain metastasis in breast cancer

Xiao-Gang Li, Doctor degree[a], Chao Niu, Master degree[a], Ping Lu, Master degree[a], Hong-Wei Wan, Master degree[a], Wen-Di Jin, Master degree[a], Chun-Xiao Wang, Master degree[a], Wen-Yuan Mao, Undergraduate degree[a], Zhi-Ping Zhang, Master degree[a], Wan-Fu Zhang, Doctor degree[a], Bo Li, Master degree[a,*]

Abstract

Background: The presence of breast cancer in the brain, also known as brain metastasis (BMS), is the primary reason for a bad prognosis in cases of breast cancer. Breast cancer is the most prevalent malignant tumor seen in women in developing nations. At present, there is no effective method to inhibit brain metastasis of breast cancer. Therefore, it is necessary to conduct a systematic study on BMS of breast cancer, which will not provide ideas and sites for follow-up studies on the treatment and inhibition of BMS.

Methods: In this study, data set GSE43837 was screened from gene expression omnibus database, and then R language tool was used for differential analysis of its expression spectrum, The gene ontology functional enrichment and Kyoto encyclopedia of genes and genomes signal pathway enrichment analyses, as well as the interactive gene retrieval tool for hub-gene analysis, were performed.

Results: According to the findings, the primary genes linked to breast cancer brain metastases are those that involve interactions between cytokines and their respective receptors and between neuroactive ligands and their respective receptors. The majority of the gene ontology enrichment took place in the extracellular structural tissues, the extracellular matrix tissues, and the second message-mediated signaling. We were able to identify 8 genes that are linked to breast cancer spreading to the brain. The gene score for matrix metallopeptidase1 (MMP-1) was the highest among them, and the genes MMP10, tumor necrosis factor alpha-inducible protein 8, collagen type I alpha 2 chain, vascular cell adhesion molecule 1, and TNF superfamily member 11 were all connected to 1 another in an interaction way.

Conclusions: There is a possibility that the 8 key genes that were identified in this research are connected to the progression of BMS in breast cancer. Among them, MMP1 is 1 that has the potential to have a role in the diagnosis and treatment of BMS in breast cancer.

Abbreviations: BBB = blood–brain barrier, BMS = brain metastasis, CC = cellular component, COL1A2 = collagen type I alpha 2 chain, GEO = gene expression omnibus, GO = gene ontology, KEGG = Kyoto encyclopedia of genes and genomes, MMP = matrix metallopeptidase, PPI = protein-protein interaction, PLCG2 = phospholipase C gamma 2, STRING = interactive gene retrieval tool, VCAM1 = vascular cell adhesion molecule 1.

Keywords: brain metastasis, breast cancer, hub-genes

1. Introduction

One of the cancers that occurs most often is breast cancer, which may be seen in women. The most frequent form of cancer in women globally is breast cancer, which, depending on the kind of breast cancer, has a metastatic disease occurrence rate of between 30% and 40%.[1] In women in the United States, breast cancer is the second major cause of brain metastases and the second greatest cause of mortality from cancer overall.[2] Patients with breast cancer who have metastases in the

X-GL and CN contributed equally to this work.

The authors have no funding and conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are publicly available.

Declarations

All Authors of this manuscript have confirmed their participation in this study, and all have also published this study in this journal. All Authors have confirmed that there is no conflict of interest. The completion of this project is funded by the Medical Reserve Talents Program of Yunnan Provincial Health Commission (H-2018064), and Authors' contributions: Xiao-gang Li, Chao Niu, and Ping Lu were responsible for the data analysis and collation of the results. Hong-wei Wang, Wen-di Jin, Chun-Xiao Wang, Wen-Yuan Mao, and Zhi-Ping Zhang were responsible for the proofreading and review of the article, while Bo LI, and Wang-fu Zhang were responsible for the foundation support and article writing proofreading

[a] Department of General Surgery, Affiliated Hospital of Yunnan University, Kunming, China.

** Correspondence: Bo Li, Department of General Surgery, Affiliated Hospital of Yunnan University, Kunming 650021, China (e-mail: libo2345442@126.com).*

How to cite this article: Li X-G, Niu C, Lu P, Wan H-W, Jin W-D, Wang C-X, Mao W-Y, Zhang Z-P, Zhang W-F, Li B. Screening and identification of hub-gene associated with brain metastasis in breast cancer. Medicine 2023;102:7(e32771).

Received: 4 October 2022 / Received in final form: 5 January 2023 / Accepted: 6 January 2023

http://dx.doi.org/10.1097/MD.0000000000032771

Key points

- This study reanalyzed the breast cancer dataset GSE43837.
- We provide 8 hub genes associated with BMS in breast cancer.
- We predict that MMP1 is a key gene for BMS in breast cancer and may serve as a potential therapeutic site.

brain are at an increased risk for both morbidity and death.[3] Patients diagnosed with brain metastasis (BMS) have a terrible prognosis, with just 10 months being the median length of survival.[4]

Breast cancer is the most frequent form of the disease seen in women in underdeveloped nations across the globe, and it is also the main cause of death from cancer.[5] Breast cancer is one of the most prevalent forms of the disease seen in urban Chinese women.[6] Cancer patients in China are responsible for 12.2% of newly diagnosed cases of breast cancer and 9.6% of breast cancer-related fatalities globally each year.[7] In China, the rate of women being diagnosed with breast cancer has been steadily increasing over the previous 2 decades. In 2015, there were a total of 268,600 persons in China who were diagnosed with breast cancer. These numbers indicate that breast cancer is the most prevalent form of cancer among Chinese women. It is anticipated that there would be 405,680 women in China who are at risk for breast cancer by the year 2025.[8–10]

Breast cancer is the leading cause of BMS, accounting for about 10% to 30% of metastatic breast cancer cases, second only to lung cancer in BMS rate.[11,12] The molecular mechanism of BMS is still unclear. Due to its very poor prognosis, many studies have focused on the tendency of circulating cancer cells to migrate to the brain in breast cancer, including the molecular mechanisms that may allow them to invade through the BLOOD-brain barrier, but there are few studies on the molecular mechanisms of BMS themselves.[13] The blood-brain barrier (BBB) is responsible for maintaining the structural integrity of the cerebrovascular system. This barrier is made up of endothelial cells and other types of cells, in addition to transport components.[14] Morad et al[4] found that breast cancer-derived exosomes can disrupt an intact blood-brain barrier. The first step of BMS is to damage the BBB to promote the invasion of tumor cells into the brain. Therefore, one of the core issues in the field of BMS is the damage degree of BBB.[15]

Although breast cancer can be detected and treated by early screening at present, most breast cancer is difficult to be diagnosed early and early, especially the molecular mechanism of brain metastasis of breast cancer is still unclear. Therefore, this study intends to analyze the molecular differences and HUB gene of primary breast cancer and breast cancer BMS through bioinformatics methods, in order to provide ideas and loci for the prevention of BMS.

2. Data sources and methods

2.1. Data sources and Identification of DEGs

The data format for MINiML was modified using the GSE43837 gene expression omnibus (GEO) query software that was obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/), and the probes that corresponded to numerous molecules were eliminated.[16] We get the data in a format known as MINiML. R Software Limma software package (version 3.40.2) was used in order to investigate differences in the expression of mRNA. The adjusted P values were analyzed in GEO to correct the false positive results. "Adjusted P & Lt; 0.05 and log2 (multiple change) & gt; 1 or log2 (multiple variation) < -1" was defined as a threshold for differential mRNA expression screening. When the probes corresponding to the same molecule were encountered, only the probe with the maximum signal value was retained. Then, the sample standardization was checked through the box diagram. PCA and UMAP [version 0.2.7.0] were used to check the clustering between sample groups. Then limMA package [version 3.42.2] was used to conduct difference analysis between the 2 groups.[17] For specific difference analysis, see difference analysis results in "Results." In this study, software R (version 3.6.3) was used for statistical analysis and visualization, and GGploT2 package [version 3.3.3] and Complex Heatmap package [version 2.2.0] were used for visualization analysis.[18]

2.2. Functional enrichment analysis of DEGs

The data were evaluated using a technique called feature enrichment so that possible capabilities of prospective targets may be identified even further. Gene ontology, often known as gene ontology (GO), is a method that is frequently used for the annotation of functional genes, in particular molecular functions, biological pathways, and cellular components. The Kyoto encyclopedia of genes and genomes (KEGG) enrichment analysis is a helpful tool for studying gene function and the information that is connected to high-level genomic function.[19] The Cluster Profiler tool in R was used to conduct an analysis of the GO function of possible mRNA and enhance the KEGG pathway.[20] This was done in order to get a deeper comprehension of the oncogenic role that target genes play. The GGplot2 software tool in R was used to generate the boxplot. The PCA diagram was created using the GGord software tool in the R programming language. The R software package known as Pheat map is responsible for presenting the term heatmap. Both the Ggplot2 package (version 3.3.3) and the cluster Profiler tool (version 3.14.3) are used for the viewing and analysis of chosen data, respectively.[18,21]

2.3. Protein-protein interaction (PPI) network construction

As a protein complex network, PPI network is formed under the action of biochemistry or electrostatic force.[22] PPI networks are very important in the functioning of molecular processes, and PPI abnormalities are often the root cause of a wide variety of illnesses, including cancer.[23] PPI networks were constructed with the use of the interactive gene retrieval tool (STRING) database (https://string-db.org/cgi/input.pl) for the purpose of this research. To add more value to each of the modules, use the FunRich tool. Eight hub genes with high connectivity were selected and located into PPI based on STRING with a confidence value of 0.4 and a maximum number of interactions of 5. Co-expression analysis of HUB genes was conducted using STRING.

3. Results

3.1. Identification of DEGs in breast cancer

The GSE43837 dataset included 19 samples of BMS and 19 samples of primary breast cancer. As shown in Figure 1A, the median of each sample is basically at the same level, which indicates a good degree of normalization among samples. After filtering, the total number of molecules is 19705, among them meet |log2(FC)|>1 & *P* value < .05, There were 963 ids at the threshold of .05, at which the number of high expression (positive logFC) was 591 in the BMS group and 372 in primary breast cancer. Meet |log2(FC)|>1.5 & *P* value < .05, There were 208 ids at the .05 threshold, at which 114 were overexpressed (log FC positive) in the BMS group and 94 in primary breast cancer. Meet |log2(FC)|>2 & *P* value < .05, There were 53 ids at the threshold of .05, at

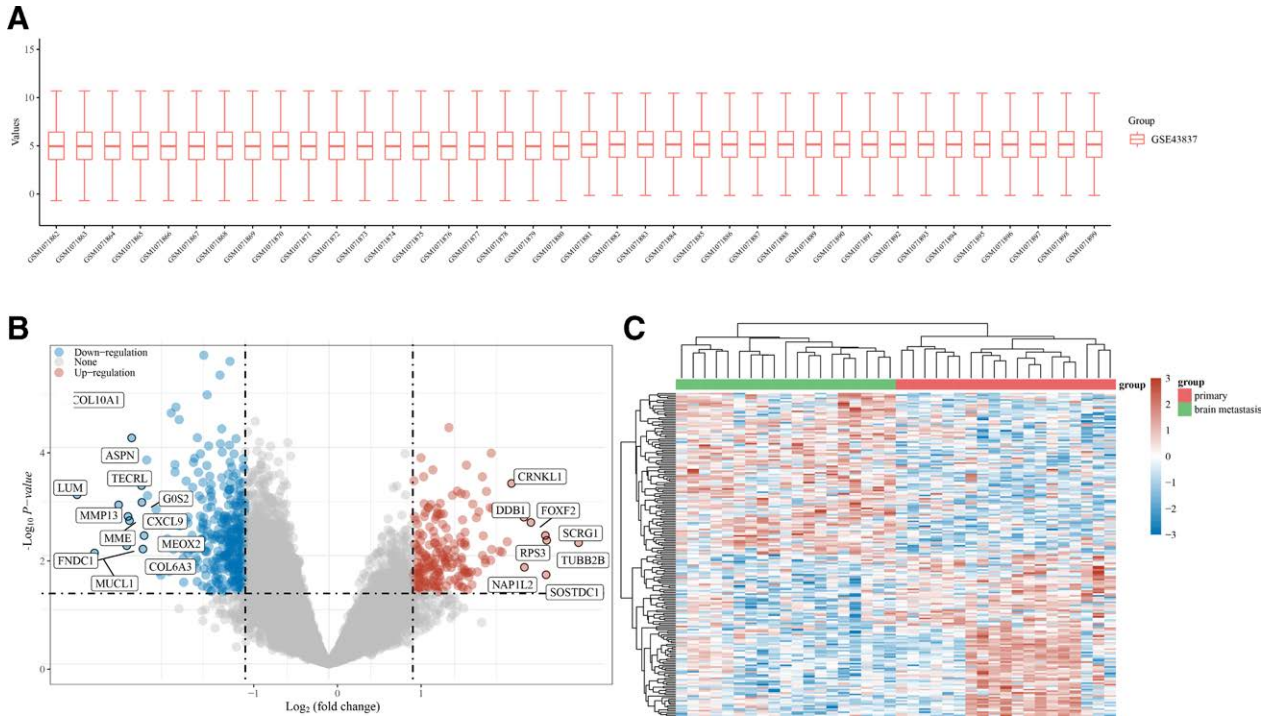

Figure 1. (A): A visualization of the data using a standardized back box. The rows indicate samples, and the colors reflect the gene expression levels found in those samples. Each data set is represented by a distinct hue. (B): A differential gene volcano map: a volcano map constructed using fold change and corrected *P* values the picture contains X dots, which indicate genes whose expression levels have dramatically increased, and X dots, which represent genes whose expression levels have significantly decreased. (C): Differential gene expression heat map: a heat map of differential gene expression, with distinct hues representing the expression trend in various tissues. As a result of the huge number of genes that exhibit differential expression, the top 50 genes that have undergone an upregulation and the top 50 genes that have undergone a downregulation are displayed below, respectively.

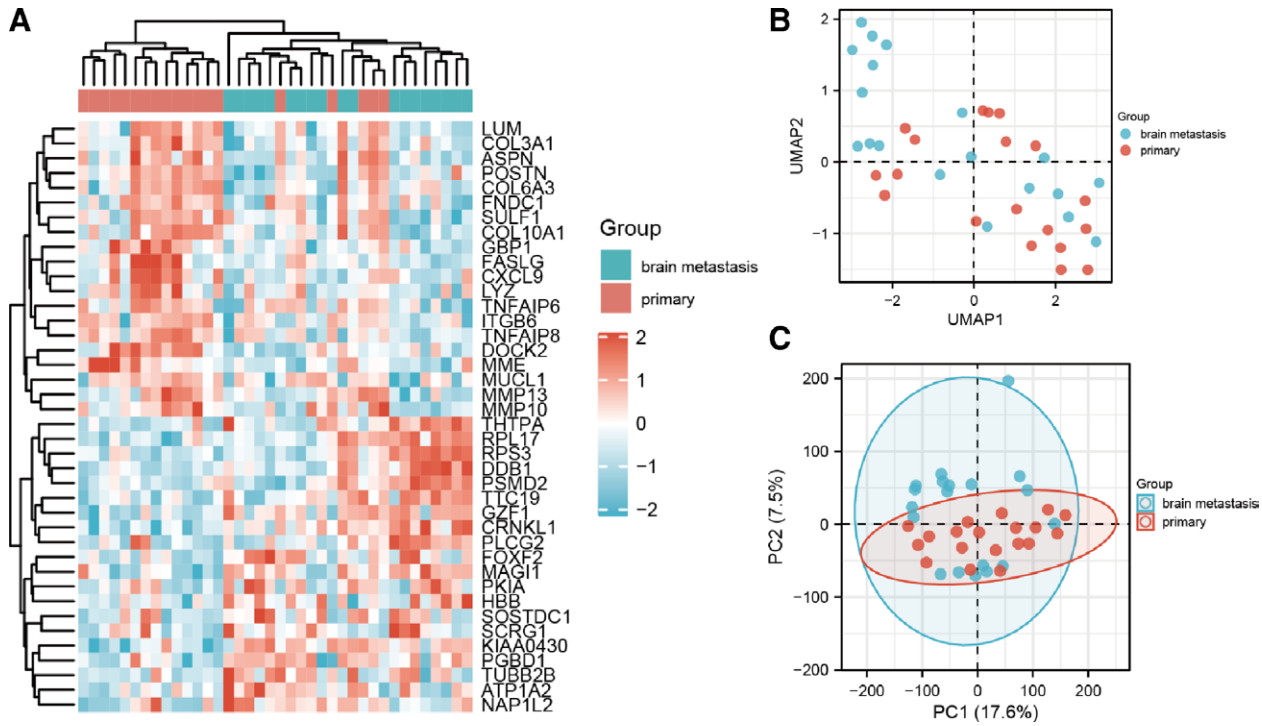

Figure 2. (A): Heatmap: Expression of top20 genes with high and low expression in the visualization expression spectrum. (B): PCA graph-sample clustering. (C): UMAP diagram - Sample clustering.

which 34 ids were overexpressed (log FC positive) in BMS and 19 ids were overexpressed (log FC negative) in primary breast cancer (Fig. 1B, C).

We chose the top 20 up-regulated and down-regulated genes for breast cancer for the purpose of reconstructing a heat map in order to do more research and discover the essential genes

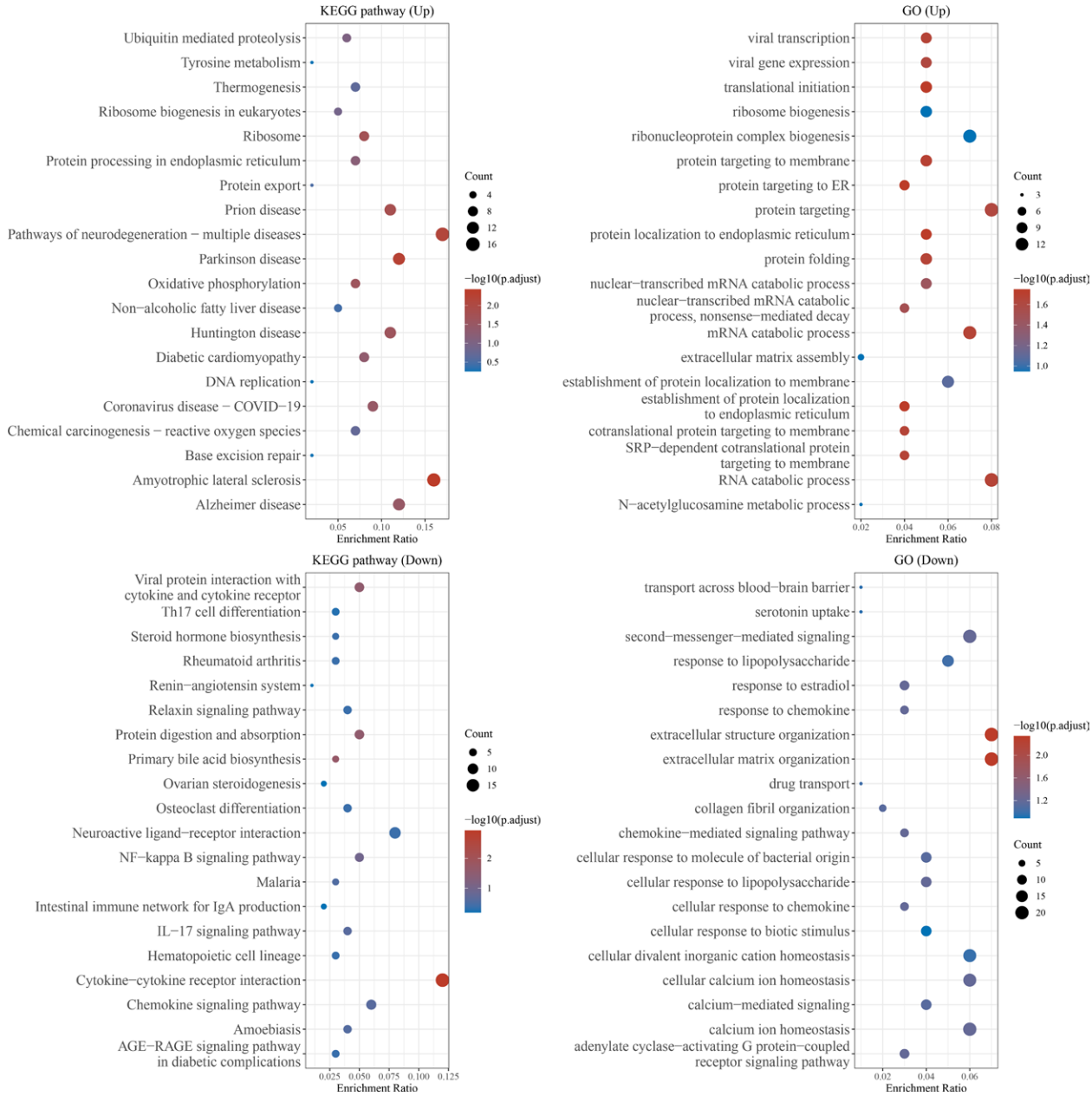

Figure 3. Functional enrichment: findings of enrichment for the KEGG pathway of genes that were differently up-regulated; Results of GO word enrichment analysis for genes that were significantly up-regulated; KEGG pathway enrichment findings of genes with differentially decreased expression levels; The GO word enrichment findings of differently down-regulated genes and the functional enrichment results were acquired using the R software program Cluster Profiler (Version: 3.18.0). The importance of differentially down-regulated genes is shown by various colors. The lower the FDR value is, the greater the value that is being referred to. The size of the circle is proportional to the number of genes that were enriched. In enrichment results, P & Lt; 0.05 or FDR & Lt; 0.05 considered to be enriched to a meaningful pathway, i.e., the right ruler of the enrichment diagram -log10 (P). GO = gene ontology, KEGG = Kyoto encyclopedia of genes and genomes.

connected to BMS of breast cancer while also removing duplicate information (Fig. 2A). PC1 (principal component 1) represents 17.6% of the differences between the 2 groups that can explain the comprehensive analysis results, and PC2 (principal component 2) represents 7.5% of the differences between the 2 groups that can explain the comprehensive analysis results. The smaller the distance between points is, the more similar it is; otherwise, the greater the difference is. It can be seen from the figure that there should be significant differences between the 2 groups. As shown in Figure 2C, there is a small difference between BMS and primary breast cancer. When the samples of each group are separated, it indicates that there are significant differences between the groups, and the subsequent difference analysis may have more meaningful results. UMAP results showed few differences between BMS and primary breast cancer (Fig. 2B)

3.2. Functional enrichment analysis of DEGs

Breast cancer has an overabundance of KEGG signaling pathways. The results of the BMS analysis indicated that the KEGG signaling pathway including downregulated genes was mostly enriched in cytokine-cytokine receptor interactions and neuroactive ligand-receptor interactions. The KeGG signaling pathway of elevated genes was mostly enriched in the pathways associated with amyotrophic lateral sclerosis and neurodegeneration-multiple

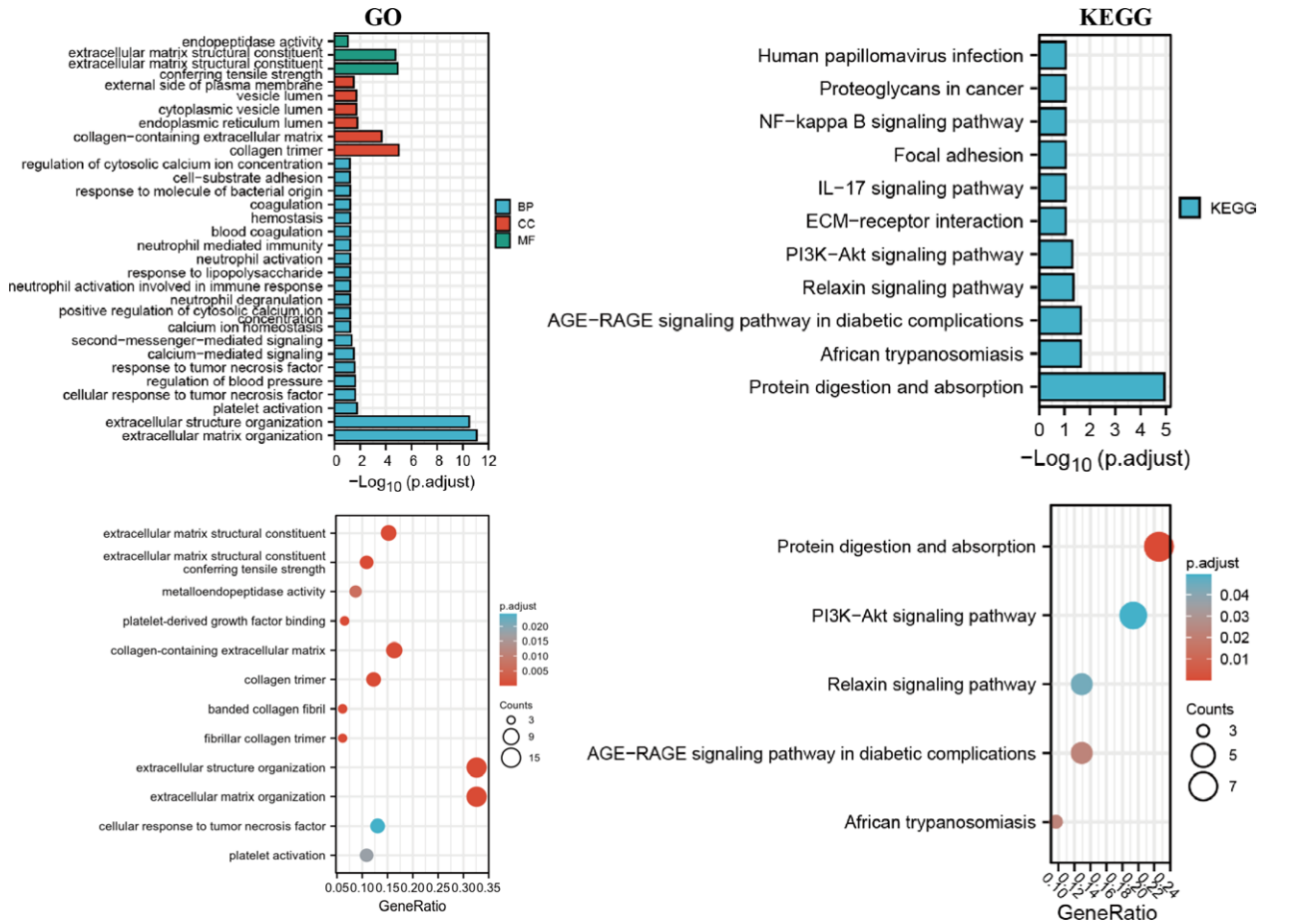

Figure 4. Functional enrichment: KEGG's findings for enrichment of pathways. Enrichment findings for GO keywords showing 20 genes with upregulation and 20 genes with downregulation. By using Cluster Profiler, we were able to acquire enrichment findings for the KEGG pathway that showed 20 up-regulated and 20 down-regulated genes (version: 3.18.0). where the importance of the differential enrichment findings is shown by various hues depending on the color. The lower the FDR number, the greater the value being measured. The size of the circle is proportional to the number of genes that were enriched. The greater the number, the wider the circumference of the circle. In enrichment results, P & Lt; 0.05 or FDR & Lt; 0.05 is considered to be enriched to a meaningful pathway, that is, the right bar of the enrichment diagram -log10 (P) is greater than 1.3. GO = gene ontology, KEGG = Kyoto encyclopedia of genes and genomes.

diseases. The results of the GO enrichment analysis performed on breast cancer brain metastases revealed the following: The GO terms for down-regulated genes were mostly enriched in ribonucleoprotein complex formation, protein targeting, mRNA catabolic processes, and RNA catabolic processes. The majority of the genes that were enriched throughout the process of upregulation were those that were involved in extracellular structural organization, extracellular matrix organization, second messenger-mediated signaling, cytosolic calcium homeostasis, and calcium ion homeostasis (Fig. 3).

The bar charts and bubble charts that were used for the GO enrichment study revealed that two of the most important biological processes, known as back propagations, are the organization of extracellular structures and extracellular matrix. Molecular Function is richer in extracellular matrix structural constituents and extracellular matrix structural constituents that impart tensile strength. The essential cellular component is an extracellular matrix that contains collagen and collagen trimer. According to the results of the Kegg signal pathway enrichment, the differential genes were mostly localized in the protein digestion and absorption route. (Fig. 4, Tables 1, 2).

3.3. Hub gene expression

We found 8 differentially expressed genes, They are matrix metallopeptidase 10 (MMP-10), MMP1, tumor necrosis factor alpha-inducible protein 8 and vascular cell adhesion molecule 1 (VCAM1), collagen type I alpha 2 chain (COL1A2), and plasma-derived growth factor receptor alpha (PDGFRA), phospholipase C gamma 2 (PLCG2) and tumor necrosis factor superfamily member 11. MMP1 is the only one of them that has a very high connection. It is possible that these genes have a direct connection to the development of BMS in breast cancer (Fig. 5).

4. Discussion

In this study, R language package was used to analyze GEO dataset GSE43837, and sequencing data of 19 patients with BMS were selected as the experimental group and 19 patients with primary breast cancer as the control group. In comparison to primary breast cancer, the level of mRNA expression in BMS was shown to be significantly different. This was determined by differential analysis. We showed via gene function analysis that the differential mRNAs of BMS were mostly connected to cell contact and cell communication. This was the case both inside and between individual cells. The neuroactive ligand receptor interaction pathway is mostly made up of a set of neuroreceptor genes. These genes take role in the processing of environmental information and the interaction of signal molecules.[24] Some examples of these genes are dopamine receptors and proto-oncogenes. According to the findings of our study, cellular signaling may entail interaction routes between neuroactive ligand

Table 1

GO enrichment results.

Ontology	ID	Description	GeneRatio	Bg Ratio	*P* value	p. adjust	Q value
BP	GO: 0030198	extracellular matrix organization	15/46	368/18670	5.83e–15	8.53e–12	7.11e–12
BP	GO: 0043062	extracellular structure organization	15/46	422/18670	4.33e–14	3.17e–11	2.64e–11
BP	GO: 0030168	platelet activation	5/46	153/18670	3.62e–05	0.018	0.015
BP	GO: 0071356	cellular response to tumor necrosis factor	6/46	291/18670	7.55e–05	0.024	0.020
CC	GO: 0005581	collagen trimer	6/49	87/19717	7.44e–08	9.82e–06	7.36e–06
CC	GO: 0005583	fibrillar collagen trimer	3/49	11/19717	2.35e–06	1.03e–04	7.74e–05
CC	GO: 0098643	banded collagen fibril	3/49	11/19717	2.35e–06	1.03e–04	7.74e–05
CC	GO: 0062023	collagen-containing extracellular matrix	8/49	406/19717	6.50e–06	2.14e–04	1.61e–04
MF	GO: 0030020	extracellular matrix structural constituent conferring tensile strength	5/46	41/17697	6.63e–08	1.20e–05	8.79e–06
MF	GO: 0005201	extracellular matrix structural constituent	7/46	163/17697	1.96e–07	1.77e–05	1.30e–05
MF	GO: 0048407	platelet-derived growth factor binding	3/46	11/17697	2.67e–06	1.61e–04	1.18e–04
MF	GO: 0004222	metalloendopeptidase activity	4/46	103/17697	1.46e–04	0.007	0.005

BP = back propagation, CC = cellular component, GO = gene ontology, MF = molecular functions.

Table 2

KEGG enrichment results.

ONTOLOGY	ID	Description	GeneRatio	BgRatio	*P* value	p. adjust	Q value
KEGG	hsa04974	Protein digestion and absorption	7/31	103/8076	9.14e–08	1.15e–05	1.01e–05
KEGG	hsa05143	African trypanosomiasis	3/31	37/8076	3.64e–04	0.023	0.020
KEGG	hsa04933	AGE-RAGE signaling pathway in diabetic complications	4/31	100/8076	5.39e–04	0.023	0.020
KEGG	hsa04926	Relaxin signaling pathway	4/31	129/8076	.001	0.044	0.039
KEGG	hsa04151	PI3K-Akt signaling pathway	6/31	354/8076	.002	0.050	0.044

KEGG = Kyoto encyclopedia of genes and genomes.

receptors and general attachment mechanisms.[25] It is interesting to observe that extracellular structures and extracellular matrix include a substantially higher concentration of differentially expressed genes. By controlling the interaction between growth factors and extracellular molecules, extracellular matrix plays a significant role in the formation of muscle cells, as well as the maintenance of structures, the transmission of forces, and the remodeling of tissue.[26] The KEGG enrichment analysis showed that the differential genes were mostly focused in the processes of protein digestion and absorption.

Through examination of the core genes, we were able to identify 8 important hinge genes. MMP10, matrix metallopeptidase (yes), tumor necrosis factor alpha-inducible protein 8, and vascular cell adhesion molecule 1 are some of the enzymes that have been linked to cardiovascular disease VCAM1. MMP1, also known as COL1A2 and plasma-derived growth factor receptor alpha (PDGFRA), PLCG2, and tumor necrosis factor superfamily member 11 (MMP1 is a matrix metalloproteinase). It has been shown that as a result of its participation in the disintegration of the extracellular matrix, it encourages the invasion of cancer cells. In addition to this, he is implicated in the early phases of the development of metastatic disease in a variety of malignancies[27,28] Research has shown that the MMP family, of which MMP1 is a part, is linked to invasion and metastasis.[29] Research has shown that MMP1 has a role in the process of tissue remodeling, as well as tumor invasion and metastasis.[29,30] Research has shown that MMP1 has a role in the process of tissue remodeling, as well as tumor invasion and metastasis. Notably, McP-1 and MMP10 are versatile effectors that are also engaged in a variety of physiological and pathological processes. These include prolonged stimulation of the endothelial cells, malfunction of the endothelial cells, and increased vascular permeability.[31] In addition to degrading ECM to promote invasion, Mmp10 is also thought to promote metastasis through proteolytic regulation of the extracellular pool of growth factors such as TGF-β and Wnt.[32] Meanwhile, we found that other parts of HUB genes (VCAM1 COL1A2, PDGFRA, PLCG2 and tumor necrosis factor superfamily member 11) were involved in invasion, metastasis and intercellular communication of tumor cells. From brain metastases from breast cancer is usually the main factors of breast cancer death, understanding the molecular mechanism of brain metastases from breast cancer or molecular markers and subsequent brain metastasis related target detection and treatment under the premise of our study is based on the depth of the data set, data mining, obtain 8 central associated with brain metastases from breast cancer gene. This may be the basis of our subsequent anti-BMS or BMS mechanism research.

Acknowledgments

TCGA and GEO belong to public databases. Our study is based on open-source data, so there are no ethical issues and other conflicts of interest.

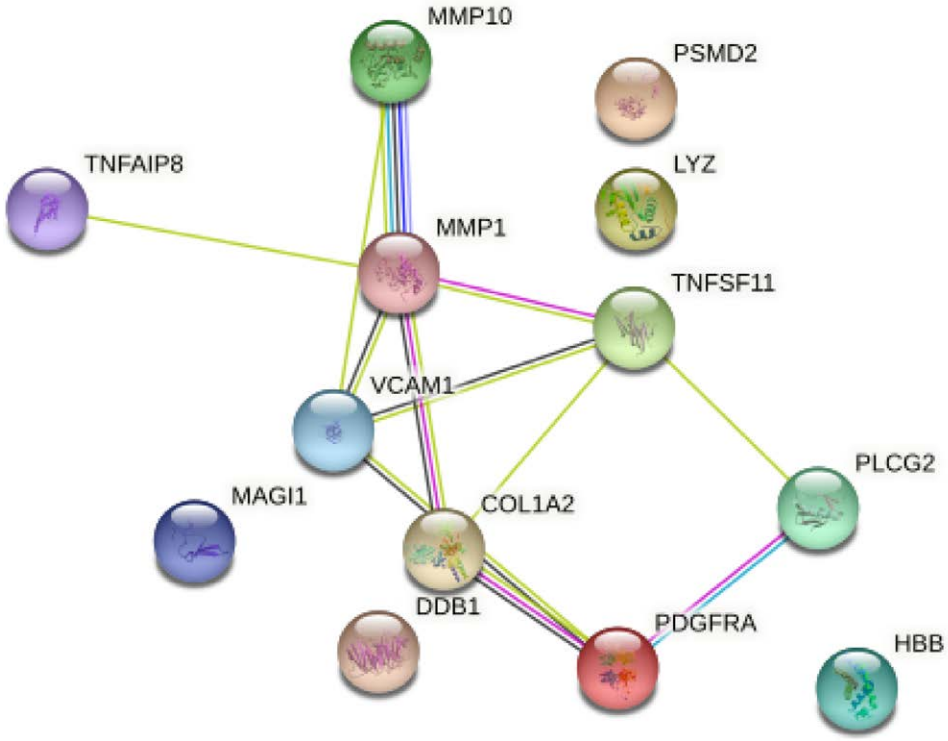

Figure 5. PPI networks of differentially expressed genes. PPI = protein-protein interaction.

Author contributions

Conceptualization: Wen-Di Jin, Wan-Fu Zhang, Bo Li.
Data curation: Xiao-Gang Li, Chao Niu, Ping Lu.
Funding acquisition: Wen-Di Jin.
Investigation: Hong-Wei Wan, Wen-Di Jin, Chun-Xiao Wang, Wen-Yuan Mao, Zhi-Ping Zhang.
Resources: Hong-Wei Wan, Chun-Xiao Wang, Wen-Yuan Mao, Bo Li.
Software: Zhi-Ping Zhang, Wan-Fu Zhang.
Supervision: Wan-Fu Zhang, Bo Li.

References

[1] Liu L, Wang Y, Miao L, et al. Combination immunotherapy of MUC1 mRNA nano-vaccine and CTLA-4 blockade effectively inhibits growth of triple negative breast cancer. Mol Ther. 2018;26:45–55.
[2] Kim JM, Miller JA, Kotecha R, et al. Stereotactic radiosurgery with concurrent HER2-directed therapy is associated with improved objective response for breast cancer brain metastasis. Neuro Oncol. 2019;21:659–68.
[3] Martin AM, Cagney DN, Catalano PJ, et al. Brain metastases in newly diagnosed breast cancer: a population-based study. JAMA Oncol. 2017;3:1069–77.
[4] Morad G, Carman CV, Hagedorn EJ, et al. Tumor-derived extracellular vesicles breach the intact blood-brain barrier via transcytosis. ACS Nano. 2019;13:13853–65.
[5] Downs BM, Mercado-Rodriguez C, Cimino-Mathews A, et al. DNA methylation markers for breast cancer detection in the developing world. Clin Cancer Res. 2019;25:6357–67.
[6] Xue D, Sun H, Li PP. Long-term Chinese herbs decoction administration for management of hot flashes associated with endocrine therapy in breast cancer patients. Chin J Cancer Res. 2011;23:74–8.
[7] Chen HL, Ding A, Wang FW. Prognostic effect analysis of molecular subtype on young breast cancer patients. Chin J Cancer Res. 2015;27:428–36.
[8] Wu Q, Liao W, Zhang M, et al. Cost-effectiveness of tucatinib in human epidermal growth factor receptor 2-positive metastatic breast cancer from the US and Chinese perspectives. Front Oncol. 2020;10:1336.
[9] Lei W, Li H, Song G, et al. Efficacy and safety of fulvestrant 500mg in hormone-receptor positive human epidermal receptor 2 negative advanced breast cancer: a real-world study in China. J Cancer. 2020;11:6612–22.
[10] Tian S, Jing R, Zhang W. Network-based approach to identify the antiproliferative mechanisms of bruceine D in breast cancer from the cancer genome atlas. Front Oncol. 2020;10:1001.
[11] Cordero A, Kanojia D, Miska J, et al. FABP7 is a key metabolic regulator in HER2+ breast cancer brain metastasis. Oncogene. 2019;38:6445–60.
[12] Sirkisoon SR, Carpenter RL, Rimkus T, et al. TGLI1 transcription factor mediates breast cancer brain metastasis via activating metastasis-initiating cancer stem cells and astrocytes in the tumor microenvironment [published correction appears in oncogene. 2021 Mar;40(12):2338]. Oncogene. 2020;39:64–78.
[13] Ebright RY, Zachariah MA, Micalizzi DS, et al. HIF1A signaling selectively supports proliferation of breast cancer in the brain. Nat Commun. 2020;11:6311.
[14] Drilon A, Siena S, Dziadziuszko R, et al. Entrectinib in ROS1 fusion-positive non-small-cell lung cancer: integrated analysis of three phase 1-2 trials [published correction appears in Lancet Oncol. 2020 Feb;21(2):e70] [published correction appears in Lancet Oncol. 2020 Jul;21(7):e341]. Lancet Oncol. 2020;21:261–70.
[15] Lyle LT, Lockman PR, Adkins CE, et al. Alterations in pericyte subpopulations are associated with elevated blood-tumor barrier permeability in experimental brain metastasis of breast cancer. Clin Cancer Res. 2016;22:5287–99.
[16] Davis S, Meltzer PS. GEOquery: a bridge between the gene expression omnibus (GEO) and BioConductor. Bioinformatics. 2007;23:1846–7.
[17] Smyth GK. Limma: linear models for microarray data. In: Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Springer, New York, NY. 2005:397–420.
[18] Gu Z, Roland E, Matthias S. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32:2847–9.
[19] Pathan M, Keerthikumar S, Ang CS, et al. FunRich: an open access standalone functional enrichment and interaction network analysis tool. Proteomics. 2015;15:2597–601.

[20] Kanehisa M, Furumichi M, Tanabe M, et al. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017;45:D353–61.
[21] Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS J Integr Biol. 2012;16:284–7.
[22] Franceschini A, Szklarczyk D, Frankild S, et al. STRING v9.1: protein–protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 2013;41:D808–15.
[23] Wong FH, Huang CY, Su LJ, et al. Combination of microarray profling and protein–protein interaction databases delineates the minimal discriminators as a metastasis network for esophageal squamous cell carcinoma. Int J Oncol. 2009;34:117–28.
[24] Ji X, Bossé Y, Landi MT, et al. Identification of susceptibility pathways for the role of chromosome 15q25.1 in modifying lung cancer risk. Nat Commun. 2018;9:3221.
[25] Castellani CA, Longchamps RJ, Sumpter JA, et al. Mitochondrial DNA copy number can influence mortality and cardiovascular disease via methylation of nuclear DNA CpGs. Genome Med. 2020;12:84.
[26] Knuiman P, Hangelbroek R, Boekschoten M, et al. Impact of protein supplementation during endurance training on changes in skeletal muscle transcriptome. BMC Genomics. 2020;21:397.
[27] Halberg N, Sengelaub CA, Navrazhina K, et al. PITPNC1 recruits RAB1B to the golgi network to drive malignant secretion. Cancer Cell. 2016;29:339–53.
[28] Au Yeung CL, Co NN, Tsuruga T, et al. Exosomal transfer of stroma-derived miR21 confers paclitaxel resistance in ovarian cancer cells through targeting APAF1. Nat Commun. 2016;7:11150.
[29] Guey LT, García-Closas M, Murta-Nascimento C, et al. Genetic susceptibility to distinct bladder cancer subphenotypes. Eur Urol. 2010;57:283–92.
[30] Pankratova EV, Stepchenko AG, Portseva T, et al. Different N-terminal isoforms of Oct-1 control expression of distinct sets of genes and their high levels in Namalwa Burkitt's lymphoma cells affect a wide range of cellular processes. Nucleic Acids Res. 2016;44:9218–30.
[31] Hsu KS, Guan BJ, Cheng X, et al. Translational control of PML contributes to TNFα-induced apoptosis of MCF7 breast cancer cells and decreased angiogenesis in HUVECs. Cell Death Differ. 2016;23:469–83.
[32] Rowbotham SP, Li F, Dost AFM, et al. H3K9 methyltransferases and demethylases control lung tumor-propagating cells and lung cancer progression. Nat Commun. 2018;9:4559.