

Screening and Identification of Key Biomarkers in Breast Cancer with Brain Metastasis: Evidence from Bioinformatic Analysis

tao ming Shao

Department of Breast and Thyroid Surgery, Jiangmen Central Hospital <https://orcid.org/0000-0002-7866-9439>

zhi yang Hu

The Affiliated Hospital of Xiangnan University

wen wei Li

Jiangmen Central Hospital

long yun Pan (✉ gdgzwcwk@163.com)

<https://orcid.org/0000-0001-9626-3718>

Research article

Keywords: breast cancer, brain metastases, bioinformatic analysis, prognosis; biomarker

Posted Date: July 13th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-38542/v1>

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

[Read Full License](#)

Abstract

Purpose. Breast cancer (BC) has a poor prognosis when brain metastases (BM) occur, and the treatment effect is limited. In this study, we aim to identify representative candidate biomarkers for clinical prognosis of patients with BM and explore the mechanisms underlying the progression of BC.

Methods. Herein, we examined the Microarray datasets (GSE125989) obtained from the Gene Expression Omnibus database to find the target genes in BC patients with BM. We employed the GEO2R tool to filter the differentially expressed genes (DEGs) that participate in primary BC and BC with BM. Subsequently, using the DAVID tool, we conducted an enrichment analysis with the screened DEGs based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO) functional annotation. The STRING database was employed to analyze the protein-protein interactions of the DEGs and visualized using Cytoscape software. Lastly, the Kaplan-Meier plotter database was employed to determine the prognostic potential of hub genes in BC.

Results. We screened out 311 upregulated DEGs and 104 downregulated DEGs. The enrichment analyses revealed that all the DEGs were enriched in the biological process of extracellular matrix organization, cell adhesion, proteolysis, collagen catabolic process and immune response. The significant enrichment pathways were focal adhesion, protein absorption and digestion, ECM-receptor interaction, PI3K-Akt signalling pathway, and Pathways in cancer. The top ten hub nodes screened out included FN1, VEGFA, COL1A1, MMP2, COL3A1, COL1A2, POSTN, DCN, BGN and LOX. The Kaplan-Meier plotter results showed that the three hub genes (FN1, VEGFA and DCN) are candidate biomarkers for clinical prognosis of patients with BM.

Conclusion. we identified seven genes related to poor prognosis in BCBM. FN1, VEGFA and DCN can be considered as potential prognostic markers for BCBM. Meantime, COL1A1, POSTN, BGN and LOX may be linked to the distant transformation of BC.

Introduction

Breast cancer is one of the most common tumours in females globally and one of the cancers that cause high mortality. About 2.1 million cases and more than 626,000 deaths were recorded in 2018[1]. Breast cancer (BC) has a poor prognosis when brain metastases (BM) occur, and the treatment effect is limited. A previous study has shown that more than half of patients with brain metastases dead of central nervous system progression [2]. Considering the high morbidity, high mortality and the limitations of effective treatment of BC patients with brain metastases, it is particularly important to conduct timely screening and intervention for patients at high-risk and develop individualized treatment. High-throughput microarrays are one of the practical tools for identifying biomarkers to diagnose and detect BC phenotype features with distant metastasis. Therefore, by analyzing the microarray gene expression profile of BC patients with brain metastases, we could establish a comprehensive PPI network to obtain biomarkers that can be used for clinical prognosis of BM patients.

Herein, we first analyzed the data on gene expression profiles retrieved from the GEO database to find DEGs in breast cancer with brain metastatic and matched non-brain metastatic breast cancer. Then, the DEGs were used for GO function annotation analysis and KEGG pathway enrichment analysis. Next, a PPI network was established to identify hub genes related to breast cancer brain metastases. Finally, through survival analysis, further search for genes related to prognosis among the important genes screened. These results will help us understand the molecular basis of breast cancer with brain metastatic and provide potential targets for future clinical intervention.

Methods

Data Retrieval

We downloaded gene expression profile of GSE125989 datasets from the GEO (<http://www.ncbi.nlm.nih.gov/geo/>) database. GSE125989 was based on Agilent GPL571 (Affymetrix Human Genome U133A 2.0 Array) and contained 16 samples each of non-metastatic primary BC and BC with BM.

Identification of DEGs

We used GEO2R in the identification of the DEGs, then statistically compared their expressions between the primary BC group and BC with the BM group. We chose the DEGs according to the standard: $p < 0.05$ and $|\logFC| > 1$.

Enrichment analysis of DEGs

Using DAVID tool [3], we performed KEGG pathway and GO functional annotation enrichment analyses[4] with the screened DEGs.

Assessment of PPI network and identification of hub genes

We employed the STRING database [5] to examine the PPI of the DEGs, which were visualized using the Cytoscape software. The Cytohubba plug-in based on Cytoscape was utilized to perform the hub gene analysis [6, 7], and the top rank 10 genes with degrees > 50 in all modules were considered to be the hub genes. Subsequently, the KEGG analyses for hub genes in this module were performed using DAVID.

Prognostic Potential of hub genes in Breast Cancers

The Kaplan-Meier (KM) plotter can be employed to evaluate the influence of up to 54,000 genes on prognosis in 21 types of cancer. Currently, breast cancer has the largest dataset ($n = 6,234$) followed by ovarian, lung, and gastric cancers with 2,190, 3,452, and 1,440 datasets, respectively. The miRNA subsystems comprise 11,000 samples sourced from 20 different types of cancer and include RNA-seq data and gene chip- data for the databases are sourced from TCGA, GEO, and EGA. Primarily, this tool is used to discover and validate prognostic markers based on meta-analysis. Herein, we employed KM plotter (<http://kmplot.com/analysis/>) to determine the correlation between the expression of hub genes

and BC prognosis [8]. Also calculated were the log-rank P-value and the hazard ratio (HR) with 95% confidence intervals.

Results

Identification Of Degs In Bc With Bm

We selected GSE125989 containing data on 16 samples each of non-metastatic primary BC and BC with BM samples for analysis of DEG through GEO2R. Subsequently, we identified 415 DEGs based on the stander: $P < 0.05$ and $|\log FC| > 1$. The DEGs included 311 upregulated and 104 downregulated genes (Fig. 1).

GO and KEGG analyses of DEGs

We performed the analyses mentioned above to explore further the main roles of the DEGs, as well as key pathways in BC patients with BM. The GO function enrichment was categorized as follows: biological process (BP), cellular component (CC), and molecular function (MF). The upregulated DEGs that were considerably enriched in MF included: calcium ion binding, heparin-binding, extracellular matrix structural constituent, serine-type endopeptidase activity and integrin-binding, et al. (Fig. 2A). For BP, enrichment of the upregulated DEGs was observed in extracellular matrix organization, cell adhesion, proteolysis, collagen catabolic process and immune response, et al. Furthermore, the CC analysis indicated that the significantly enriched were in extracellular region, extracellular exosome, extracellular space, extracellular matrix, and proteinaceous extracellular matrix, et al. The signalling pathway analysis by KEGG show that the upregulated DEGs were enriched in Focal adhesion, ECM-receptor interaction, PI3K-Akt signalling pathway, Protein absorption and digestion, Proteoglycans in cancer, Pathways in cancer, and et al. (Table 1). On the other hand, in BP, enrichment of the downregulated DEGs were mainly observed in ubiquitin-dependent protein catabolic process, positive regulation of the apoptotic process, chemical synaptic transmission, canonical glycolysis, and glycolytic process, et al. (Fig. 2B). For MF, they were mainly enriched in the following activities: protein binding, thiol-dependent ubiquitin-specific protease, translation initiation factor and guanyl-nucleotide exchange factor, SUMO binding and phosphatidylinositol-4,5-bisphosphate 5-phosphatase. Additionally, the CC terms for these genes were related to the cytoplasm, cytosol, membrane, nucleoplasm and perinuclear region of cytoplasm, et al. The KEGG pathway results showed that the down-regulated DEGs were enriched in the HIF-1 signalling pathway, Insulin secretion, Glycolysis / Gluconeogenesis, Fructose and mannose metabolism, Carbon metabolism, and Aldosterone-regulated sodium reabsorption (Table 2).

Table 1
KEGG pathway enrichment analysis of up DEGs. P < 0.05

Term	Count	PValue	Genes
Focal adhesion	24	1.40E-11	COMP, CAV1, COL1A1, COL1A2, COL3A1, COL5A1, COL5A2, COL5A3, COL6A1, COL6A2, COL6A3, CCND2, FN1, FLNA, IGF1, LAMA4, MYL9, MYL9P, PDGFD, PDGFRA, PDGFRB, THBS1, THBS2, THBS4
PI3K-Akt signaling pathway	22	4.85E-06	COMP2R, COL1A1, COL1A2, COL3A1, COL5A1, COL5A2, COL5A3, COL6A1, COL6A2, COL6A3, CCND2, FN1, IGF1, IL2RA, LAMA4, PDGFD, PDGFRA, PDGFRB, THBS1, THBS2, THBS4
ECM-receptor interaction	16	1.36E-10	COMP, COL1A1, COL1A2, COL3A1, COL5A1, COL5A2, COL5A3, COL6A1, COL6A2, COL6A3, FN1, HSPG2, LAMA4, THBS1, THBS2, THBS4
Protein digestion and absorption	15	1.76E-09	COL1A1, COL1A2, COL3A1, COL5A1, COL5A2, COL5A3, COL6A1, COL6A2, COL6A3, COL10A1, COL14A1, COL15A1, DPP4, ELN, MME
Pathways in cancer	15	0.0269901	CXCL12, RUNX1T1, F2R, EDNRA, FN1, FOXO1, IGF1, LAMA4, MMP2, PDGFRA, PDGFRB, PTGER3, PTGER4, RARB, STAT1
Proteoglycans in cancer	13	6.86E-04	TIMP3, CAV1, DCN, ESR1, FN1, FLNA, HSPG2, IGF1, IL12B, LUM, MMP2, PLA1, THBS1
Cell adhesion molecules (CAMs)	12	1.28E-04	CD40LG, CD6, CDH15, CLDN17, HLA-G, HLA-DRA, PECAM1, PTPRC, SELP, SIGLEC1, VCAN
HTLV-I infection	11	0.0324955	CD3E, CD3G, CCND2, IL2RA, KAT5, HLA-G, HLA-DRA, NFATC4, PDGFRA, PDGFRB
Amoebiasis	9	0.0012954	COL1A1, COL1A2, COL3A1, COL5A1, COL5A2, COL5A3, FN1, IL12B, LAMA4
Phagosome	9	0.0107756	COMP, C1R, HLA-G, HLA-DQB1, HLA-DRA, MRC2, THBS1, THBS2, THBS4

Table 2
KEGG pathway enrichment analysis of down DEGs. P < 0.05

Term	Count	P Value	Genes
HIF-1 signaling pathway	6	4.60E-04	CAMK2G, CUL2, EIF4E, HK2, PRKCA, VEGFA
Insulin secretion	5	0.002585	ATP1A2, CAMK2G, PCLO, PRKCA, SNAP25
Glycolysis / Gluconeogenesis	4	0.01055	ALDOB, HK2, PGK1, TPI1
Carbon metabolism	4	0.041508	ALDOB, HK2, PGK1, TPI1
Fructose and mannose metabolism	3	0.019896	ALDOB, HK2, TPI1
Aldosterone-regulated sodium reabsorption	3	0.02884	ATP1A2, NEDD4L, PRCKA

Construction of PPI network and Hub genes analysis

The STRING database was applied to construct the PPI network to explore the molecular mechanisms about BC with BM progression. In total, 390 nodes and 2,197 edges were identified among the DEGs (Fig. 3A). Subsequently, Cytoscape software was employed for further assessment of the network to enable us to understand the DEGs with regards to the PPI network. Further, the hub genes were selected with select degree > 50 as the standard by the plug-in CytoHubba, and we extracted the top ten hub nodes with higher degrees. They included FN1, VEGFA, COL1A1, MMP2, COL3A1, COL1A2, POSTN, DCN, BGN and LOX (Fig. 3B). Moreover, the DAVID online tool was used to predict possible enrichment pathways of the hub genes, and the pathways closely related to Focal adhesion, ECM-receptor interaction, Amoebiasis, PI3K-Akt signalling pathway, Proteoglycans in cancer, Protein absorption and digestion, Platelet activation, Bladder cancer and Pathways in cancer (Table 3).

Table 3
KEGG pathway enrichment re-analysis of the ten hub genes. $P < 0.05$

Term	Count	P Value	Genes
Focal adhesion	5	1.12E-05	COL3A1, VEGFA, COL1A2, COL1A1, FN1
ECM-receptor interaction	4	3.80E-05	COL3A1, COL1A2, COL1A1, FN1
Amoebiasis	4	6.88E-05	COL3A1, COL1A2, COL1A1, FN1
PI3K-Akt signaling pathway	5	8.61E-05	COL3A1, VEGFA, COL1A2, COL1A1, FN1
Proteoglycans in cancer	4	4.54E-04	VEGFA, DCN, MMP2, FN1
Protein digestion and absorption	3	0.002347	COL3A1, COL1A2, COL1A1
Platelet activation	3	0.005058	COL3A1, COL1A2, COL1A1
Bladder cancer	2	0.035245	VEGFA, MMP2
Pathways in cancer	3	0.041894	VEGFA, MMP2, FN1

Prognostic Potential of hub genes in Breast Cancer

To examine the prognostic potential of the ten hub genes in Breast Cancer, KM plotter database was employed to determine the hub genes prognostic value based on Affymetrix microarrays. As shown in Fig. 4A, the three hub genes (FN1, VEGFA and DCN) were correlated with worse OS (FN1: HR = 1.28, 95% CI = 1.03 to 1.59, $P = 0.023$; VEGFA: HR = 1.35, 95% CI = 1.09 to 1.68, $P = 0.0055$ and DCN: HR = 0.69, 95% CI = 0.56 to 0.86, $P = 0.00071$) and in Fig. 4B, the seven hub genes (FN1, VEGFA, COL1A1, POSTN, DCN, BGN and LOX) were correlated with worse DMFS (FN1: HR = 1.28, 95% CI = 1.03 to 1.59, $P = 0.023$; VEGFA: HR = 1.35, 95% CI = 1.09 to 1.68, $P = 0.0055$; COL1A1: HR = 1.53, 95% CI = 1.09 to 2.14, $P = 0.012$; POSTN: HR = 1.61, 95% CI = 1.16 to 2.24, $P = 0.0042$, DCN: HR = 0.82, 95% CI = 0.67 to 0.99, $P = 0.042$; BGN: HR = 1.29, 95% CI = 1.07 to 1.57, $P = 0.0092$; and LOX: HR = 1.28, 95% CI = 1.06 to 1.55, $P = 0.012$). These results suggest the FN1, VEGFA and DCN genes affect breast cancer prognosis.

Discussion

Evidence indicates that breast cancer patients with brain metastasis (BM) have a very poor prognosis and a short survival time following BM [9]. However, much progress has been made regarding the diagnosis of the condition, especially with the application of PET-CT and MRI [10]. Despite these, the progress towards effective therapies has been slow, mainly because the standard treatment options for BM, such as surgical operation, chemotherapy, and radiotherapy are ineffective [11]. As such, there is a need to identify sensitive markers to improve the prognosis of BC with BM.

First, we screened the differential expressed genes which may be associated with BM in BC in our study. GO term analysis revealed that most of the DEGs participate in tumorigenesis through regulating cell

proliferation, differentiation, apoptosis adhesion, and migration, as well as angiogenesis. Also, the KEGG pathway analysis indicated that most of the DEGs had a strong relation to the development of tumours through the key signalling pathways, including the PI3K-AKT pathway, HIF-1 signalling pathway, pathways in cancer and the metabolic pathway [12–14]. Subsequently, the PPI network was constructed to explore the interactions of the DEGs. The top ten hub genes were significantly related to the pathway of Focal adhesion, ECM-receptor interaction, Amoebiasis, PI3K-Akt signalling pathway, Proteoglycans in cancer, Protein absorption and digestion, Platelet activation, Bladder cancer and Pathways in cancer (Table 1). FN1 and VEGFA were the most connected nodes in the module, in which FN1 (fibronectin 1 protein) regulates the migratory and adhesion ability of cells. Upregulation of FN1 has been observed BC metastases [15]. VEGFA has a vital and dual function in angiogenesis and tumour metastasis and has been implicated in the pathogenesis of breast cancer [16]. Lastly, we performed survival analysis to determine the prognostic value of ten hub genes, we found that the three hub genes (FN1, VEGFA and DCN) were correlated with worse OS (Fig. 4A), and the seven hub genes (FN1, VEGFA, COL1A1, POSTN, DCN, BGN and LOX) were correlated with worse DMFS (Fig. 4B). Of the 10 genes, five genes (FN1, VEGFA, COL1A1, BGN and LOX) have been implicated in tumour progression, as well as in predicting disease outcomes. For example, COL1A1 expression increases the invasion and metastatic capacity of gastric cancer cells [17]. BGN up-regulation has been demonstrated in many malignancies [18]. Also, LOX is upregulated in many cancers, such as BC, nasopharyngeal carcinoma, hepatocellular carcinoma, head and neck tumours, colorectal cancer and is related to poor prognosis[19]. These indicating that our big data analysis above has prognostic values in the BC cohort of GEO database. Moreover, POSTN and DCN genes were identified herein for the first time as possible markers for BC prognosis.

Herein, we identified seven possible biomarkers FN1, VEGFA, COL1A1, POSTN, DCN, BGN and LOX. However, their role in the formation and progression of cancers should be investigated further. Several studies have explored the molecular dynamics that occur during the formation and development of breast cancer[20, 21]. As a result, many BC prognostic markers have been identified. Most of these studies were conducted using animal models, in vitro cell experiments, and clinical tumour samples involving small sample size[22–24]. However, the complex interplay of BC with BM and the delicate nature of the condition requires that we conduct comprehensive studies that involve large sample sizes. Luckily, substantial progress has been made so far in bioinformatics, including the establishment of high-throughput tumour databases like GEO and TCGA[25]. The databases collect and manage data which are made public for use by researchers and has so far assisted in big data analysis consisting of large-scale cohorts of BC with BM [26, 27].

The current study is unique because, other than focusing on the mechanisms by which the activation of tumour intrinsic gene affect BC pathogenesis, we focused on gene signatures in BC with BM and their effect on BC prognosis. The findings of the present study could offer new insights into the understanding of the complex interactions between BC and BC with BM. However, this study had a few limitations. Firstly, the clinical data in the GEO database was not complete, which will prevent us from performing a comprehensive survival analysis. Secondly, we only studied one data set, which may lead to biases in our

results. Thirdly, our results still need further verification. For example, Western blot and real-time PCR can be employed to verify the selected hub genes expression in the basic experiments.

Conclusions

In conclusion, we identify seven genes with poor prognosis from the BC with BM cohort of GEO database. The three genes, namely, FN1, VEGFA and DCN can be perceived as candidate genes for the prognostic prediction of BC with BM. At the same time, COL1A1, POSTN, BGN and LOX may be related to the distant transformation of breast cancer. These genes need further research to help us fully understand the relationship between them and the prognosis of breast cancer with brain metastasis. Our BM-related gene mining strategy once again validated the feasibility of using big data analysis in the discovery of prognostic value biomarkers for malignant tumors.

Abbreviations

BC

breast cancer; BM:brain metastases. DEGs:differentially expressed genes; KEGG:Kyoto Encyclopedia of Genes and Genomes; GO:Gene Ontology. BP:biological process; CC:cellular component; MF:molecular function; TCGA:The Cancer Genome Atlas; PPI:protein protein interaction.

Declarations

Acknowledgements

Not applicable

Authors' contributions

tao ming Shao and zhi yang Hu contributed equally to this work. wen wei Li and long yun Pan are co-corresponding authors. tao ming Shao and zhi yang Hu: conception and design of the work. tao ming Shao, wen wei Li, and long yun Pan: acquisition and interpretation of the data. zhi yang Hu and wen wei Li: statistical analysis of the data. tao ming Shao and long yun Pan: writing of the draft. The authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (81472849), the Guangdong Natural Science Research (2014A030313383), and Jiangmen Science and Technology Plan Project (2019A088).

Availability of data and materials

The datasets used during the current study are available from the public database.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Department of Breast and thyroid Surgery, Jiangmen Central Hospital, Jiangmen, Guangdong 529000, China

² Department of General Surgery, The First Affiliated Hospital of Jinan University, Guangzhou, Guangdong 510632, China

³ Department of Gastrointestinal Surgery, The Affiliated Hospital of Xiangnan University, Chenzhou, Hunan 423000, China

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68:394–424.
2. Bendell JC, Domchek SM, Burstein HJ, Harris L, Younger J, Kuter I, et al. Central nervous system metastases in women who receive trastuzumab-based therapy for metastatic breast carcinoma. *Cancer.* 2003;97:2972–7.
3. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2009;37:1–13.
4. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 2017;45:D353-D61.
5. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019;47:D607-D13.
6. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol.* 2014;8(Suppl 4):11.
7. Wang H, Liu J, Li J, Zang D, Wang X, Chen Y, et al. Identification of gene modules and hub genes in colon adenocarcinoma associated with pathological stage based on WGCNA analysis. *Cancer Genet.*

- 2020;242:1–7.
8. Lanczky A, Nagy A, Bottai G, Munkacsy G, Szabo A, Santarpia L, et al. miRpower: a web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients. *Breast Cancer Res Treat.* 2016;160:439–46.
 9. Yamamura J, Masuda N, Yasojima H, Mizutani M, Kuriyama K, Nakamori S, et al. Clinicopathological Factors Related to the Prognosis of Metastatic Breast Cancer Patients after Development of Brain Metastasis. *Breast Care (Basel).* 2015;10:387–92.
 10. Gaedcke J, Traub F, Milde S, Wilkens L, Stan A, Ostertag H, et al. Predominance of the basal type and HER-2/neu type in brain metastasis from breast cancer. *Mod Pathol.* 2007;20:864–70.
 11. Niwinska A, Murawska M, Pogoda K. Breast cancer subtypes and response to systemic treatment after whole-brain radiotherapy in patients with brain metastases. *Cancer.* 2010;116:4238–47.
 12. Nagpal N, Ahmad HM, Chameettachal S, Sundar D, Ghosh S, Kulshreshtha R. HIF-inducible miR-191 promotes migration in breast cancer through complex regulation of TGFbeta-signaling in hypoxic microenvironment. *Sci Rep.* 2015;5:9650.
 13. Wei C, Li HZ, Wang YH, Peng X, Shao HJ, Li HX, et al. Exogenous spermine inhibits the proliferation of human pulmonary artery smooth muscle cells caused by chemically-induced hypoxia via the suppression of the ERK1/2- and PI3K/AKT-associated pathways. *Int J Mol Med.* 2016;37:39–46.
 14. Chu A, Liu J, Yuan Y, Gong Y. Comprehensive Analysis of Aberrantly Expressed ceRNA network in gastric cancer with and without H.pylori infection. *J Cancer.* 2019;10:853–63.
 15. Loganathan J, Jiang J, Smith A, Jedinak A, Thyagarajan-Sahu A, Sandusky GE, et al. The mushroom *Ganoderma lucidum* suppresses breast-to-lung cancer metastasis through the inhibition of pro-invasive genes. *Int J Oncol.* 2014;44:2009–15.
 16. Razak NA, Abu N, Ho WY, Zamberi NR, Tan SW, Alitheen NB, et al. Cytotoxicity of eupatorin in MCF-7 and MDA-MB-231 human breast cancer cells via cell cycle arrest, anti-angiogenesis and induction of apoptosis. *Sci Rep.* 2019;9:1514.
 17. Sun H. Identification of key genes associated with gastric cancer based on DNA microarray data. *Oncol Lett.* 2016;11:525–30.
 18. Tang Z, Yang Y, Zhang J, Fu W, Lin Y, Su G, et al. Quantitative Proteomic Analysis and Evaluation of the Potential Prognostic Biomarkers in Cholangiocarcinoma. *J Cancer.* 2019;10:3985–99.
 19. Kalikawe R, Baba Y, Nomoto D, Okadome K, Miyake K, Eto K, et al. Lysyl oxidase impacts disease outcomes and correlates with global DNA hypomethylation in esophageal cancer. *Cancer Sci.* 2019;110:3727–37.
 20. Shao M-T, Hu Y-Z, Ding H, Wu Q, Pan J-H, Zhao X-X, et al. The overexpression of ZWINT in integrated bioinformatics analysis forecasts poor prognosis in breast cancer. *Translational Cancer Research.* 2020;9:187–93.
 21. Tsang JYS, Tse GM. Molecular Classification of Breast Cancer. *Adv Anat Pathol.* 2020;27:27–35.

22. Lehtinen L, Vainio P, Wikman H, Huhtala H, Mueller V, Kallioniemi A, et al. PLA2G7 associates with hormone receptor negativity in clinical breast cancer samples and regulates epithelial-mesenchymal transition in cultured breast cancer cells. *J Pathol Clin Res.* 2017;3:123–38.
23. Wong CH, Ma BBY, Hui CWC, Lo KW, Hui EP, Chan ATC. Preclinical evaluation of ribociclib and its synergistic effect in combination with alpelisib in non-keratinizing nasopharyngeal carcinoma. *Sci Rep.* 2018;8:8010.
24. Masjedi S, Zwiebel LJ, Giorgio TD. Olfactory receptor gene abundance in invasive breast carcinoma. *Sci Rep.* 2019;9:13736.
25. Padilla-Rodriguez M, Parker SS, Adams DG, Westerling T, Puleo JI, Watson AW, et al. The actin cytoskeletal architecture of estrogen receptor positive breast cancer cells suppresses invasion. *Nat Commun.* 2018;9:2980.
26. Guo T, Ma H, Zhou Y. Bioinformatics analysis of microarray data to identify the candidate biomarkers of lung adenocarcinoma. *PeerJ.* 2019;7:e7313.
27. Lu X, Gao C, Liu C, Zhuang J, Su P, Li H, et al. Identification of the key pathways and genes involved in HER2-positive breast cancer with brain metastasis. *Pathol Res Pract.* 2019;215:152475.
28. legends.

Figures

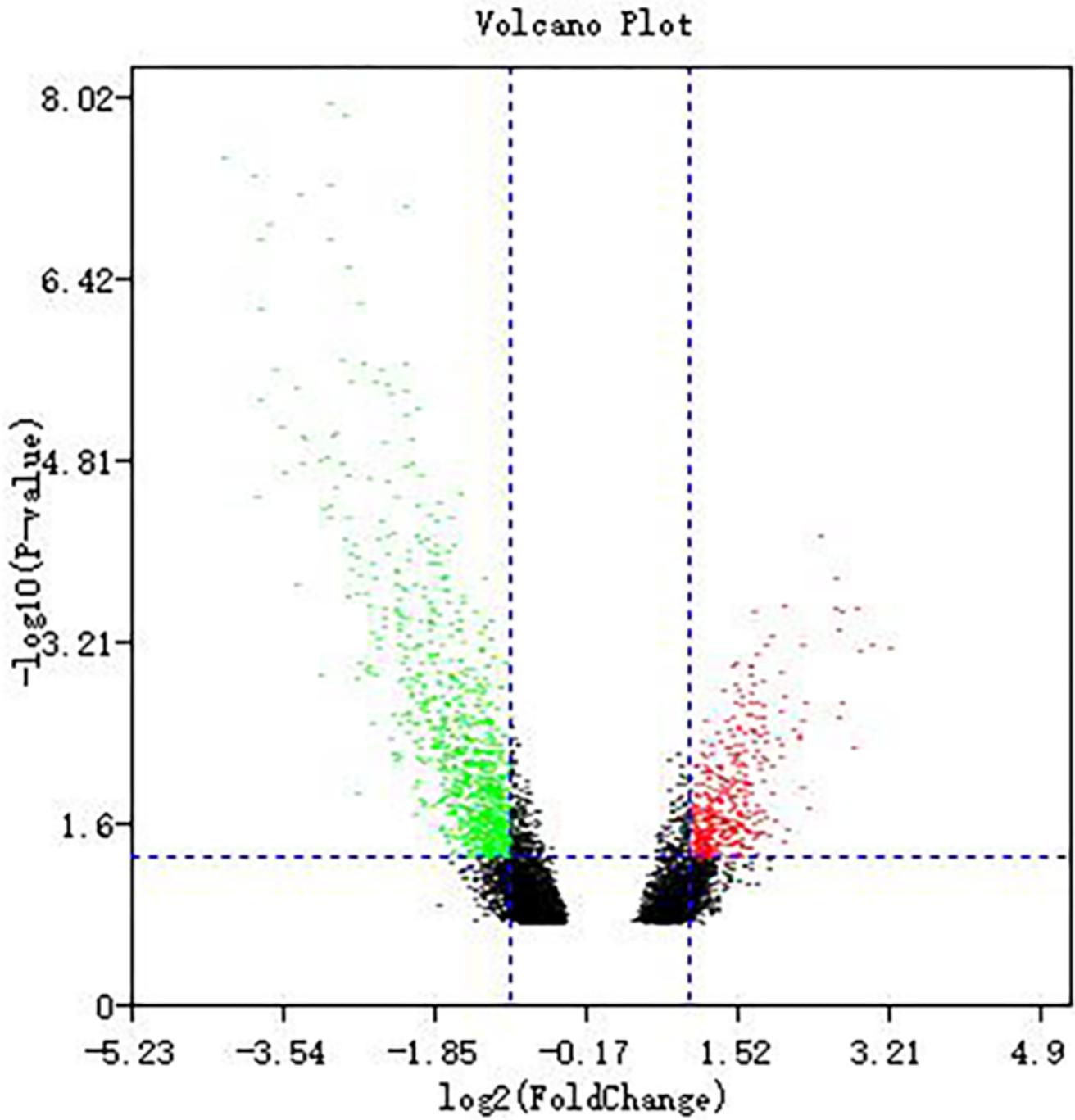


Figure 1

Volcano plot. Different colors represent different expression levels, green: down-regulated, red: up-regulated. $|\log_{2}(\text{FC})| > 1$, $P < 0.05$.

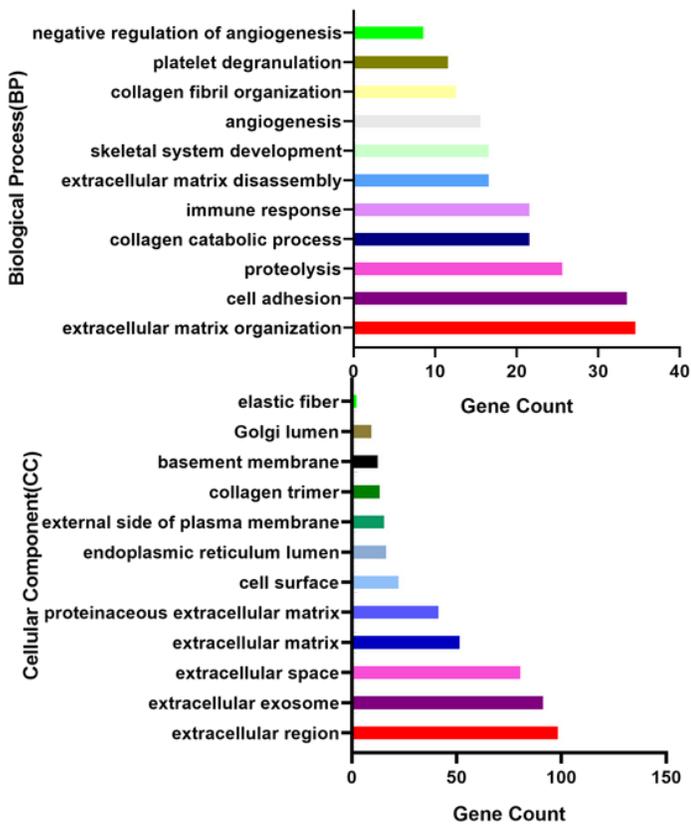


Figure 2

Gene ontology analysis of up DEGs, $P < 0.05$.

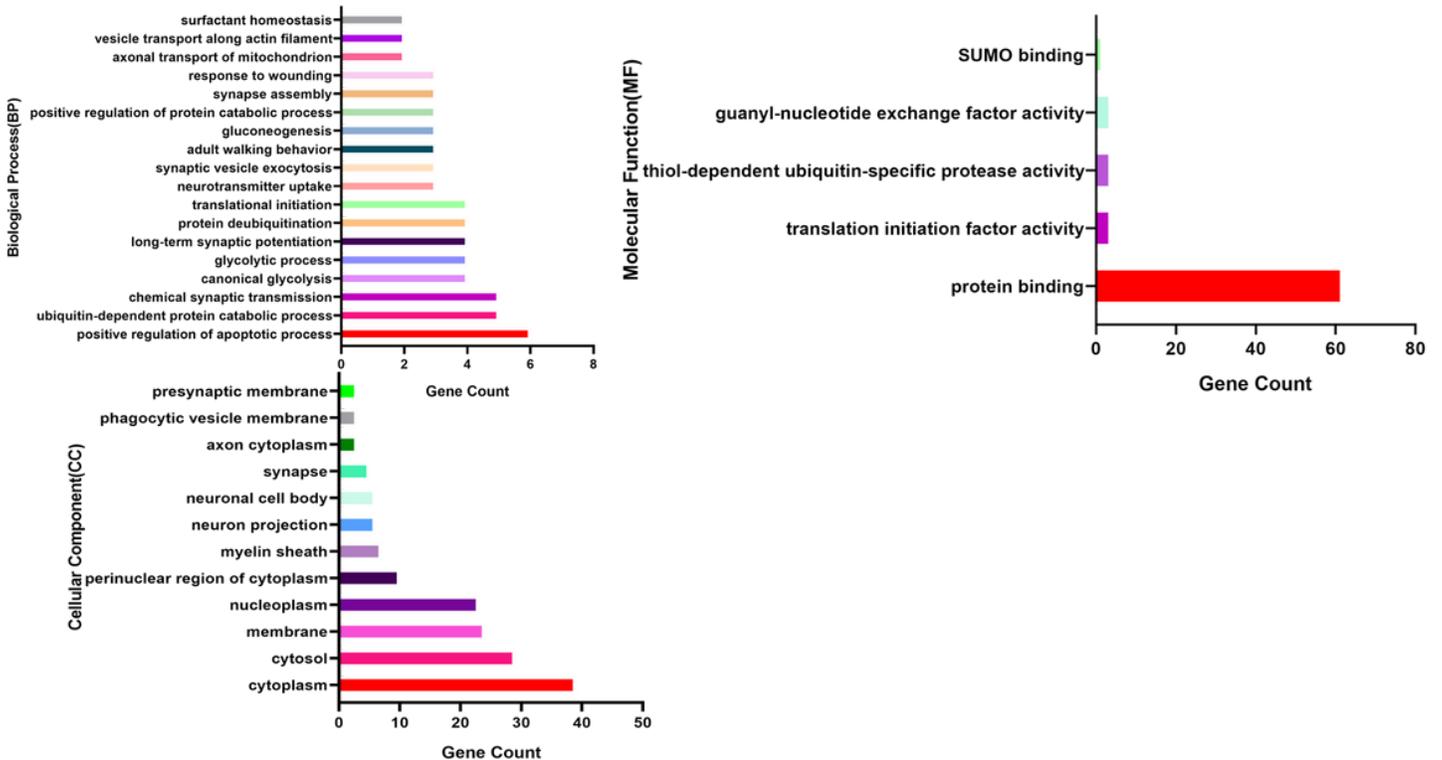


Figure 3

Gene ontology analysis of down DEGs, $P < 0.05$.

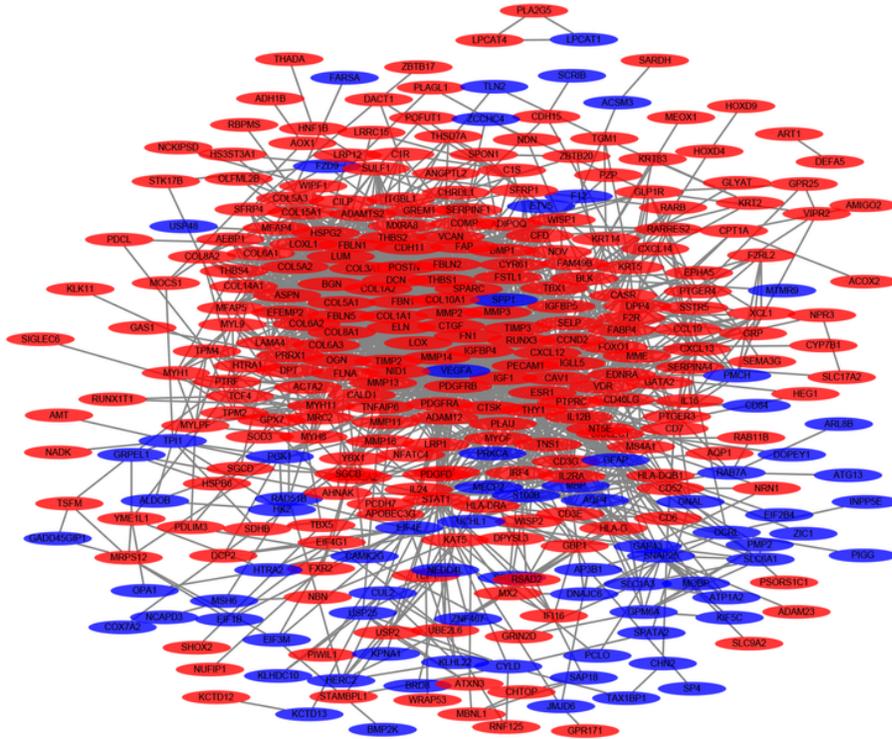


Figure 4

Protein-protein interaction network of the differentially expressed genes. The nodes indicate proteins; the edges indicate the interaction between proteins.

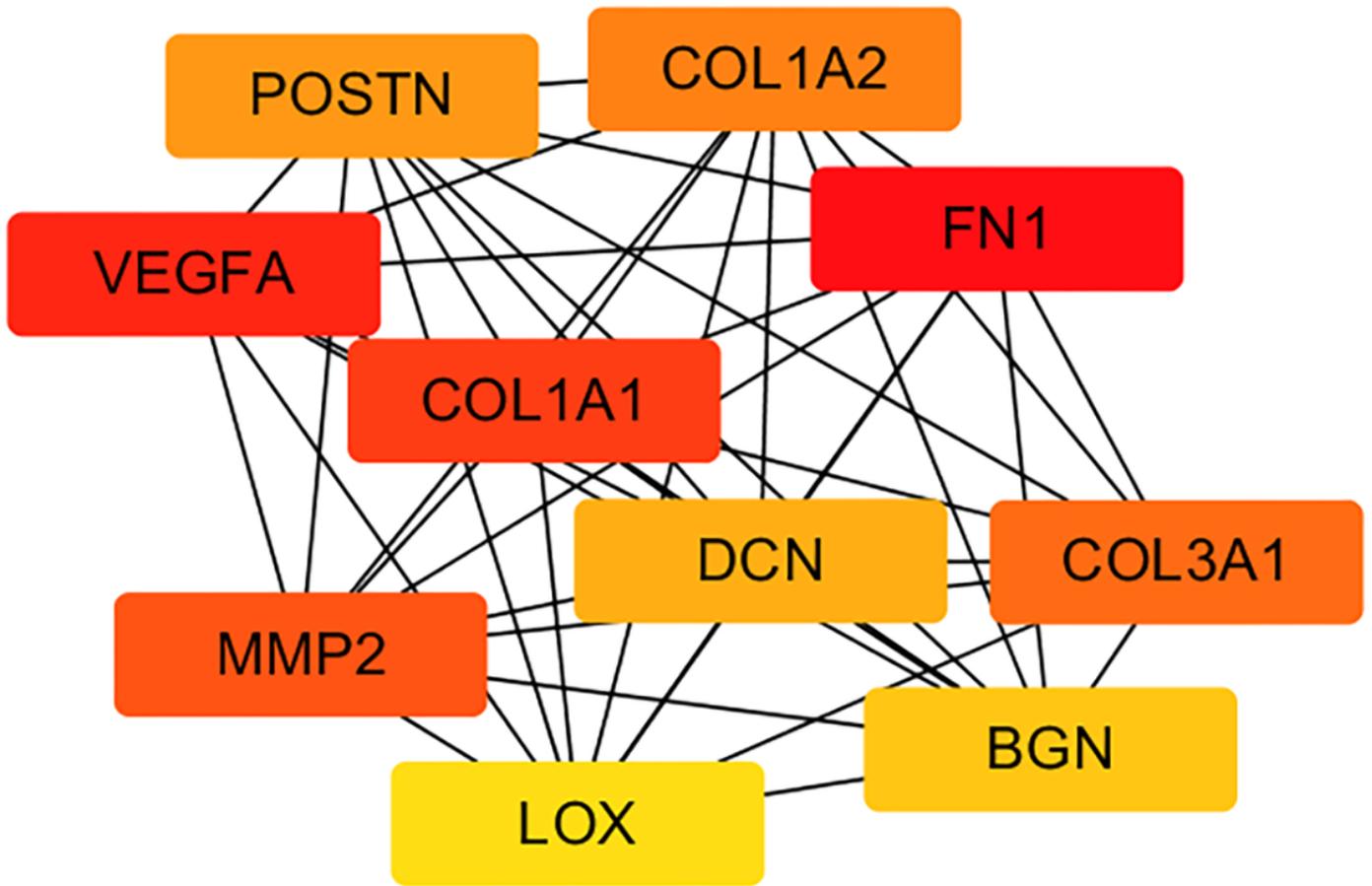


Figure 5

The top ten hub nodes with higher degrees, degree > 50.

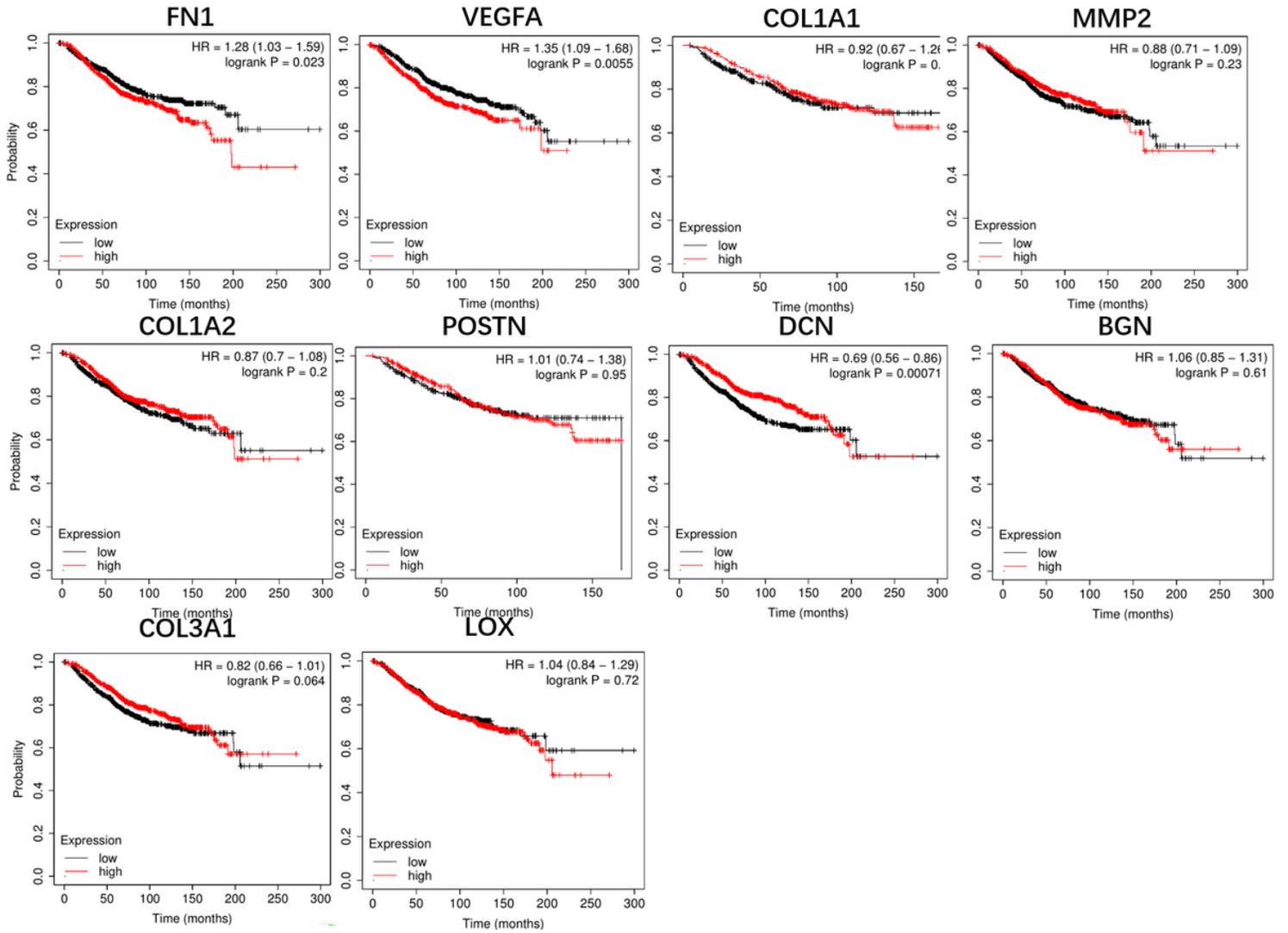


Figure 6

Prognostic information (OS) of the ten hub genes. Red: high expression, black: low expression. $P < 0.05$.

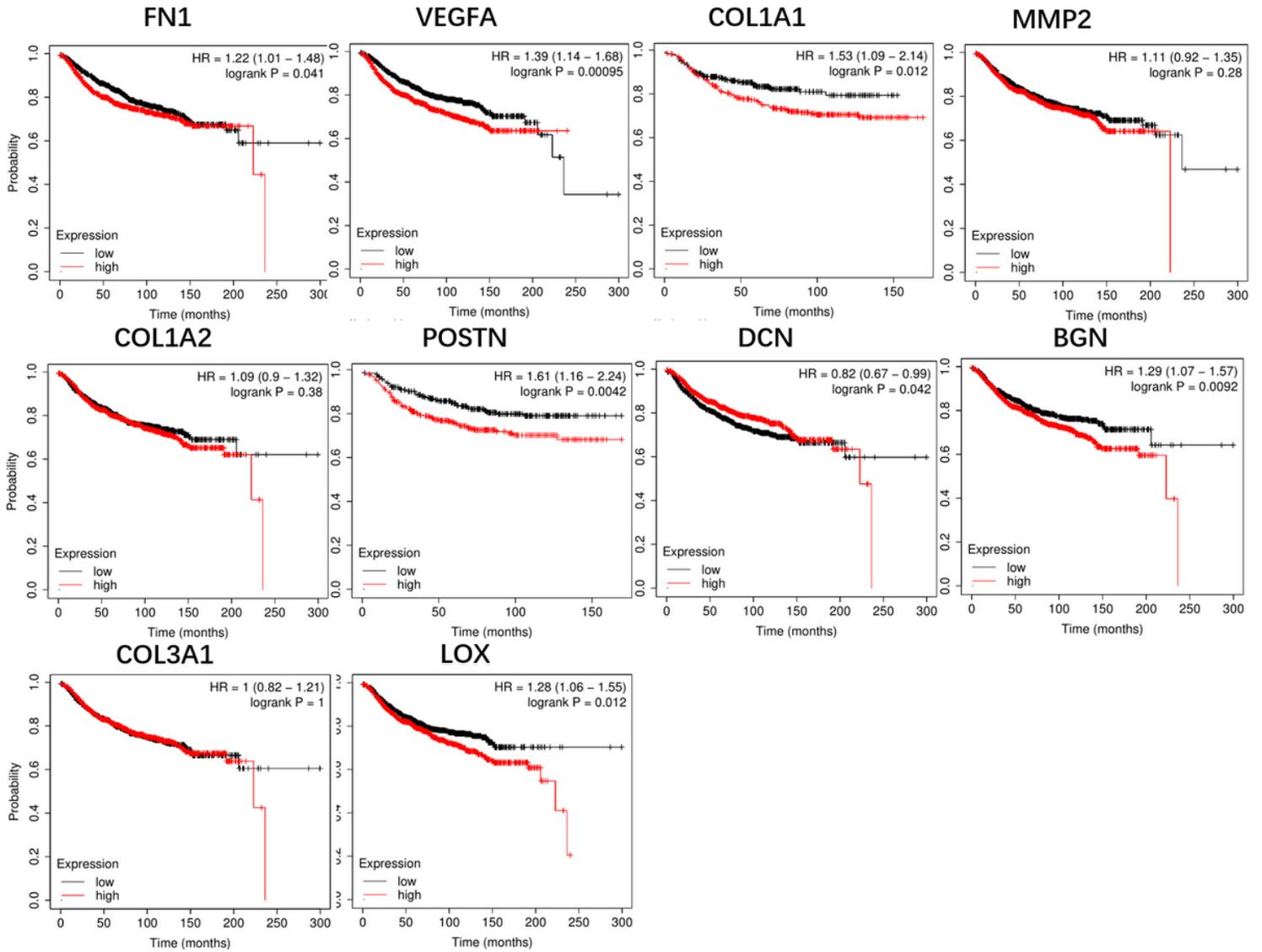


Figure 7

Prognostic information (DMFS) of the ten hub genes. Red: high expression, black: low expression. $P < 0.05$.