Candidate genes and pathways associated with brain metastasis from lung cancer compared with lymph node metastasis

XUELIAN ZHAO[1*], NAN WANG[1*], TUNGAMIRAI CHIDANGURO[1], HUANYU GU[1], YI LI[1], HUIRU CAO[1], PUSHUAI WEN[1,2] and FU REN[2]

[1]Department of Pathophysiology and [2]Biological Anthropology Institute, Jinzhou Medical University, Jinzhou, Liaoning 121001, P.R. China

Abstract. Brain metastasis from lung cancer (BMLC) is one of the common types of metastasis associated with poor prognosis. The aim of the present study was to elucidate the underlying molecular mechanisms of BMLC. The mRNA microarray dataset GSE18549 was downloaded from the Gene Expression Omnibus database. The Limma package of R was used to screen the differentially expressed genes (DEGs). Based on the DAVID database, functional and pathway enrichment analyses of DEGs were performed. Furthermore, the protein-protein interaction (PPI) network was predicted using the STRING database and visualized with Cytoscape software. In addition, hub genes and significant modules were selected based on the network. A total of 190 DEGs with $\log_2$|(fold change)|>1, including 129 significantly downregulated DEGs and 61 upregulated DEGs, were obtained. Gene Ontology functional enrichment analysis indicated that downregulated DEGs were mainly associated with 'immune response', 'cell activation' and 'leukocyte activation', while the upregulated DEGs were involved in 'DNA repair' and 'viral process'. Kyoto Encyclopedia of Genes and Genomes pathway analysis indicated that the downregulated DEGs were mainly enriched in 'chemokine signaling pathway', whereas the upregulated DEGs were associated with 'oocyte meiosis'. Based on the PPI network, 9 hub genes were selected, namely tumor necrosis factor, C-C motif chemokine ligand (CCL) 2, CD34, vascular cell adhesion molecule 1, CD48, CD27, CCL19, C-X-C motif chemokine receptor 6 and C-C motif chemokine receptor 2. The present study sheds light on the molecular mechanisms of BMLC and may provide molecular targets and diagnostic biomarkers for BMLC.

Introduction

Lung cancer is the most common cause of cancer-associated death (1). Every year, ~234,030 patients are newly diagnosed with cancer in the United States, while nearly ~154,050 patients succumb to lung cancer, with the majority of mortalities resulting from metastasis (1,2).

The brain is one of the most common sites of distant metastasis of lung cancer, particularly in lung cancer cases without lymph node involvement (3,4). The incidence of brain metastasis from lung cancer (BMLC) has been reported as 23-65% (5). The prognosis of patients with BMLC is poor, with a median survival of only 4-5 months (6). Over the past decades, several genetic alterations associated with the occurrence of BMLC were identified. For instance, vascular endothelial growth factor (VEGF)-C, a member of the VEGF family (7), was associated with BMLC. Chen *et al* (8) reported that high expression of VEGF-C is positively associated with brain metastasis in patients with lung cancer. Furthermore, the incidence of BMLC in a VEGF-C-positive expression group was higher compared with that in a negative expression group. In addition, inflammatory chemokines have been identified to be associated with brain metastases. The expression of C-X-C motif chemokine receptor (CXCR)4, receptor of the CXC chemokine ligand 12 (CXCL12), in the primary tumor tissues and distant metastatic lung tumors in the brain was reported to be higher than that in lung cancer patients without distant metastasis (9). In addition, epidermal growth factor receptor (EGFR) mutations and anaplastic lymphoma kinase (ALK) rearrangement are also thought to be associated with the BMLC (10).

Previous studies that focused on the role of genes in the occurrence of BMLC provide information on the molecular mechanisms of BMLC; however, they have remained to be fully elucidated. Recently developed gene expression profiling arrays may be used to assess the expression of thousands of genes simultaneously, providing a tool to comprehensively elucidate the mechanisms of BMLC (11).

Correspondence to: Professor Pushuai Wen, Department of Pathophysiology, Jinzhou Medical University, 40 Section III Songpo Road, Linghe, Jinzhou, Liaoning 121001, P.R. China
E-mail: wenpushuai@163.com

Professor Fu Ren, Biological Anthropology Institute, Jinzhou Medical University, 40 Section III Songpo Road, Linghe, Jinzhou, Liaoning 121001, P.R. China
E-mail: renfu1970@hotmail.com

*Contributed equally

Traditionally, lymph node metastasis was considered to be closely associated with spread of tumor cells as the origin of distant metastasis (12). Recently, a study on human colorectal cancer suggested that for most patients, metastasis to lymph nodes and distant sites were of independent origin, which suggested two different lineage associations between lymphatic and distant metastases in colorectal cancer (13). Consistent with this result, the brain was the most frequent site of distant metastasis in patients with lung cancer without lymph node metastasis (4).

In the present study, based on the data of a gene expression chip, the mRNA expression levels were compared between lung cancer with brain metastasis and lung cancer with lymph node metastasis. In addition, pathways and functional annotation was used to determine associations among the differentially expressed genes (DEGs). Furthermore, protein-protein interaction analysis was used to determine modules and hub genes. The results of the present study may enhance the current understanding of the mechanisms of BMLC.

Materials and methods

Expression profile microarray. The dataset of GSE18549, downloaded from Gene Expression Omnibus (GEO) database (www.ncbi.nlm.nih.gov/geo), is based on the GPL570 platform (Affymetrix Human Genome U133 Plus 2.0 Array). GSE18549 is a dataset containing three lymph node metastases and six brain metastases from lung cancer.

Identification of DEGs. The GSE18549 dataset was divided into two groups, namely the lymph node metastasis group and brain metastasis group. R (version 3.4.4) was used to identify the DEGs. First, background correction and normalization of the microarray data was performed in R and the Limma package (14) was then used to identify the DEGs. Multiple t-tests were used to detect the DEGs with the cut-off criteria of $\log_2$|fold change|>1 and adjusted P-value <0.05.

Gene ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. GO analysis is a common method to annotate genes and contains three categories: Cellular component (CC), molecular function (MF) and biological process (BP) (15). KEGG analysis may be used to determine the pathways of DEGs between two groups (16). GO and KEGG analyses were performed for the identified DEGs using the Database for Annotation, Visualization and Integrated Discovery (DAVID, version 6.7; http://david.abcc.ncifcrf.gov) (17). P<0.05 and the number of involved genes of ≥2 were selected as cut-off criteria.

Construction of protein-protein interaction (PPI) network and module analysis. The Search Tool for the Retrieval of Interacting Genes (STRING; version 10.5) (18) is an online tool for determining interactions among DEGs. The screened DEGs were mapped using STRING, and a combined score of >0.4 was considered to indicate significance. Cytoscape software (version 3.6.1) was then used to visualize the PPI network (19), and the Network Analyzer (version 2.7), a Cytoscape plugin, was used to compute the basic properties of the PPI network, including average clustering coefficient distribution, closeness centrality, node degree distribution and shortest path length distribution. In addition, module analysis was performed by the plug-in Molecular Complex Detection (MCODE; version 1.5.1) with the following settings: Degree cutoff, 2; node score cutoff, 0.2; k-core, 2; and maximum depth, 100, and the following criteria: MCODE score >4; number of nodes >5. Finally, the hub genes in the PPI network were determined, defined as those with a degree of connectivity of >10.

Results

Identification of DEGs. First, the raw data of GSE18549 were normalized, as presented in Fig. 1A and B. Subsequently, the significant DEGs in BMLC compared with lung cancer with lymph node metastasis were identified. A total of 190 DEGs were identified, which consisted of 129 downregulated genes and 61 upregulated genes. A volcano plot of the differential expression analysis is presented in Fig. 1C. In addition, the top 40 DEGs were displayed in a heat map in Fig. 1D.

GO terms and KEGG pathway enrichment analysis. To gain insight into the GO categories of DEGs between the lymph node metastasis group and brain metastasis group, all DEGs were uploaded to the DAVID database. The results suggested that downregulated DEGs were mainly enriched in the category BP, including 'immune response', 'cell activation' and 'leukocyte activation', while upregulated DEGs were significantly enriched in 'cell division', 'DNA repair' and 'viral process'. In the category CC, the downregulated DEGs were significantly enriched in 'plasma membrane', 'integral to plasma membrane' and 'intrinsic to plasma membrane', whereas the upregulated DEGs were significantly enriched in 'nucleoplasm', 'cytoplasm' and 'condensed nuclear chromosome'. Furthermore, in the category MF, the downregulated genes were mainly enriched in 'cytokine binding', 'nucleotide receptor activity' and 'purinergic nucleotide receptor activity', and conversely, upregulated genes were significantly enriched in 'poly-(a)-RNA binding', 'chromatin binding' and 'ATP-dependent helicase activity'. Further details on the results of the GO enrichment analysis are provided in Table I.

As presented in Table II, the KEGG pathway enrichment analysis of DEGs indicated that the downregulated DEGs were mainly enriched in 'cytokine-cytokine receptor interaction', 'natural killer cell-mediated cytotoxicity', 'hematopoietic cell lineage', 'chemokine signaling pathway' and 't-cell receptor signaling pathway', and upregulated DEGs were mainly enriched in 'oocyte meiosis'.

Construction of PPI network and module analysis. The STRING database was used to predict the interaction between the 190 DEGs, and the PPI was visualized using Cytoscape software. Initially, basic properties of the network were computed. The network contained 90 nodes and 205 interaction edges, where the average degree of connectivity (i.e., average number of neighbors) was 4.56 (Fig. 2A). Subsequently, the Network Analyzer tool was used to compute the basic properties of the PPI network, including the degree distribution, clustering coefficient, average shortest path and the closeness centrality of the PPI network (Fig. 3). The analysis indicated that the degree distribution of PPI network nodes followed the power-law
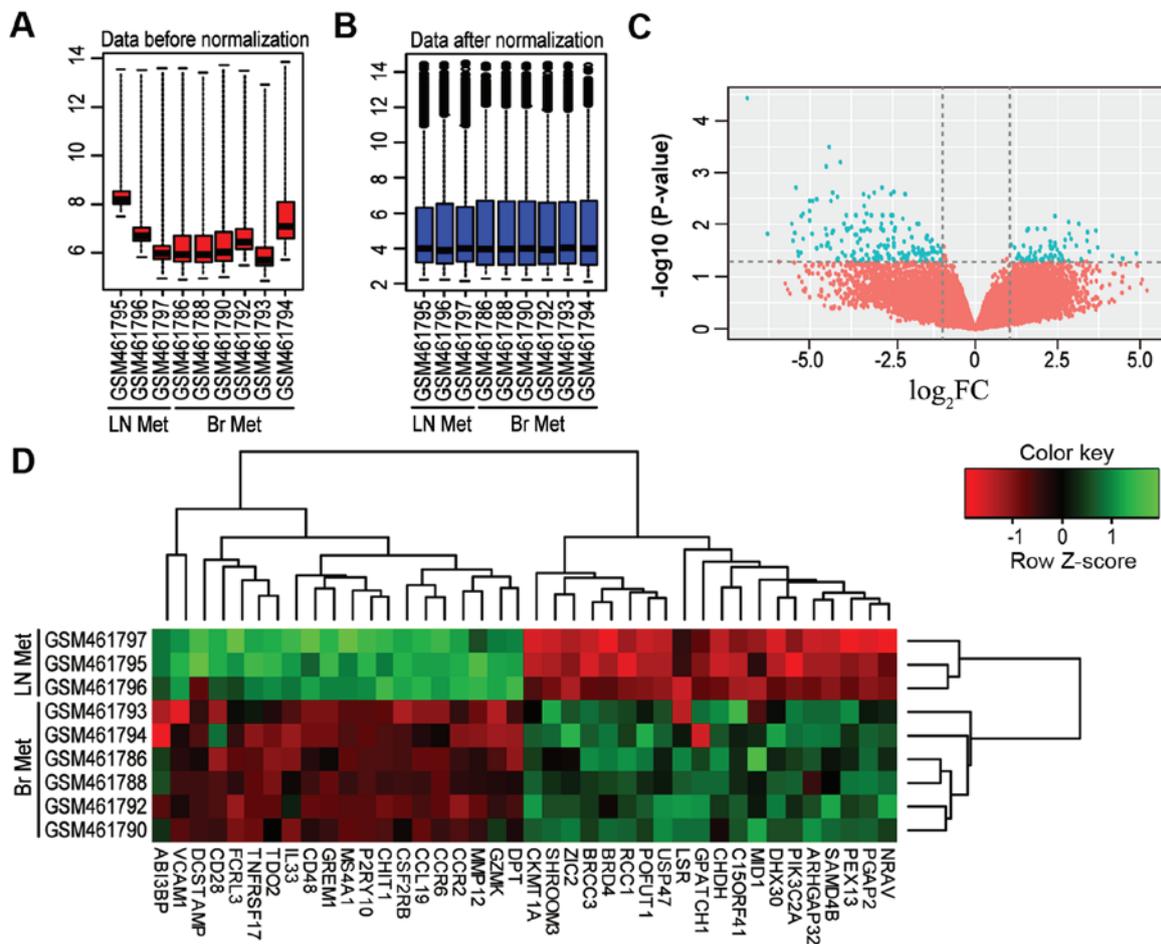
Figure 1. Normalization of gene expression data, volcano plot of DEGs and heatmap of Top 40 DEGs in mRNA expression profiling datasets. (A and B) Gene expression profiles (A) prior to normalization and (B) after normalization. (C) Volcano plot of DEGs between lymph node metastasis group and brain metastasis group. DEGs were selected with P<0.05 and log2|FC|>1. The x-axis presents the log2|FC| in gene expression between the lymph node metastasis group and brain metastasis group, and the y-axis displays the statistical significance of the differences. Red dots indicate genes without significantly different expression and blue dots indicate significant DEGs. (D) Representative heat map of the top 40 DEGs in the dataset GSE18549. Each column represents one of 40 genes and each row represents one sample. Green and red color represents high and low expression level, respectively. The color scale on the right represents the raw Z-score ranging from red to green (low to high expression level). FC, fold change; LN Met, lymph node metastasis; Br Met, brain metastasis; DEG, differentially expressed gene.

of network, and had small-world network characteristics, including a short average shortest path (20,21). In addition, one significant module consisting of 8 nodes and 28 edges was obtained from the PPI network of DEGs using MCODE (Fig. 2B). As provided in Table III, enrichment analysis suggested that the genes in this significant module were mainly associated with functional terms in the category BP, including 'G-protein-coupled receptor protein signaling pathway', 'cell surface receptor-linked signal transduction', 'taxis', 'chemotaxis' and 'locomotory behavior'. In the category CC, the genes in this significant module were significantly enriched in 'plasma membrane', and in the category MF, the genes were mainly enriched in 'C-C chemokine receptor activity', 'C-C chemokine binding', 'chemokine receptor activity', 'chemokine binding' and 'nucleotide receptor activity, G-protein coupled'. Furthermore, results from KEGG analysis demonstrated that the genes in this significant module were associated with 'chemokine signaling pathway' and 'cytokine-cytokine receptor interaction' (Table IV). Finally, the hub genes with a degree of connectivity of >10 were identified, including tumor necrosis factor (TNF), C-C motif chemokine ligand 2 (CCL2),

CD34, vascular cell adhesion molecule 1 (VCAM1), CD48, CD27, C-C motif chemokine ligand 19 (CCL19), C-X-C motif chemokine receptor 6 (CXCR6) and C-C motif chemokine receptor 2 (CCR2).

Discussion

Brain metastasis is a frequent complication in patients suffering from advanced lung cancer (3). It has been estimated that 50% of the patients diagnosed with lung cancer will develop metastatic brain lesions, which results in a dismal prognosis (22). Recently, a series of biomarkers have been identified to be associated with the development of brain metastasis, such as integrins, cell adhesion molecules, cadherins, VEGF, chemokines, matrix metalloproteinase, EGFR mutations and ALK translocations (23). However, the biology of brain metastasis is still poorly understood, as a result, the specific and effective strategies used to control or treat for brain metastasis are currently unavailable.

In the present study, a microarray dataset was analyzed to screen the DEGs between lung cancers with lymph node

Table I. GO analysis of differentially expressed genes in the GSE18549 dataset.

A, Downregulated genes

Category/GO term	P-value	Genes
Biological Process		
Immune response	2.52×10^{-20}	ITGAL, TNF, CCL2, TRGC2, FASLG, CXCL6, TLR6, CLEC10A, CHIT1, FCRL4, CD96, SH2D1A, LILRA4, POU2F2, MS4A1, LTB, CD27, CD28, CR1, CRTAM, SIT1, GBP5, CR2, PTGER4, CMKLR1, CCL19, TNFRSF17, CD180, IGSF6, CYBB, CCR6, IL18BP, KCNJ8, CCR2, LCP2
Cell activation	1.38×10^{-12}	ITGAL, CRTAM, TNF, CD3G, IKZF1, PLEK, IL21R, TLR6, SLAMF1, GIMAP1, CD48, P2RY12, VCAM1, P2RX7, MS4A1, IRF4, LTB, CD28, LCP2
Leukocyte activation	1.54×10^{-9}	ITGAL, CRTAM, CD3G, IKZF1, IL21R, TLR6, SLAMF1, GIMAP1, CD48, VCAM1, P2RX7, MS4A1, IRF4, LCP2, CD28
Defense response	8.83×10^{-9}	ITGAL, CR1, TNF, CCL2, CR2, CYSLTR1, CCL19, CXCL6, TLR6, CD180, CD48, CD84, SH2D1A, CYBB, P2RX7, CCR6, TRAC, KCNJ8, CCR2, PLA2G2D, AOC3
Lymphocyte activation	1.56×10^{-8}	CD48, VCAM1, ITGAL, P2RX7, CRTAM, CD3G, IKZF1, IL21R, MS4A1, IRF4, SLAMF1, CD28, GIMAP1
Cellular Component		
Plasma membrane	4.92×10^{-11}	TRGC2, FASLG, TLR6, DDR2, FCRL4, FCRL3, CD48, ART4, CD96, TRAC, SLC2A3, CXCR6, MS4A1, CSF2RB, RECK, CRTAM, CD3G, SIT1, GBP5, PTGER4, GPR18, CMKLR1, TNFRSF17, THY1, PRKCB, IGSF6, CD84, ARHGAP31, CCR6, CD34, CCR2, PTGDR, ADAM12, AOC3, ITGAL, CD244, TNF, CYSLTR1, CD247, BRSK1, CLEC10A, CSMD2, VCAM1, CD27, CD28, IL2RB, CR1, CR2, PLEK, SLAMF1, CD180, P2RY12, P2RY13, P2RY10, P2RX7, CYBB, LYVE1, KCNJ8, GFRA1
Integral to plasma membrane	1.14×10^{-8}	ITGAL, TNF, CYSLTR1, TRGC2, FASLG, TLR6, DDR2, CD48, CD96, TRAC, CXCR6, MS4A1, CSF2RB, CD27, CD28, IL2RB, CR1, SIT1, CD3G, CMKLR1, THY1, IGSF6, CD84, CYBB, P2RX7, LYVE1, CCR6, KCNJ8, CCR2
Intrinsic to plasma membrane	1.86×10^{-8}	ITGAL, TNF, CYSLTR1, TRGC2, FASLG, TLR6, DDR2, CD48, CD96, TRAC, CXCR6, MS4A1, CSF2RB, CD27, CD28, IL2RB, CR1, SIT1, CD3G, CMKLR1, THY1, IGSF6, CD84, CYBB, P2RX7, LYVE1, CCR6, KCNJ8, CCR2
Plasma membrane part	2.11×10^{-7}	ITGAL, CD244, TNF, CYSLTR1, CD247, TRGC2, FASLG, BRSK1, TLR6, DDR2, CD48, VCAM1, CD96, TRAC, CXCR6, MS4A1, CSF2RB, CD27, CD28, IL2RB, CR1, SIT1, CD3G, GBP5, PLEK, CMKLR1, SLAMF1, THY1, CD84, IGSF6, ARHGAP31, CYBB, P2RX7, LYVE1, CCR6, CD34, KCNJ8, CCR2
Intrinsic to membrane	3.72×10^{-7}	IL21R, TRGC2, FASLG, TLR6, DDR2, FCRL4, FCRL3, CD48, ART4, CD96, TRAC, SLC2A3, LILRA4, CXCR6, MS4A1, CSF2RB, MCOLN2, LTB, RECK, CRTAM, CD3G, SIT1, PTGER4, GPR18, CMKLR1, TNFRSF17, PKHD1L1, THY1, MCTP1, IGSF6, CD84, CCR6, CD34, PTGDR, CCR2, ADAM12, AOC3, ITGAL, CD244, TNF, CYSLTR1, CD247, KMO, CLEC10A, CSMD2, VCAM1, FMO2, CD27, CD28, IL2RB, CR1, CR2, SLAMF1, CD180, GIMAP1, P2RY12, P2RY13, CYBB, P2RY10, P2RX7, LYVE1, RNF150, KCNJ8, ST8SIA4, GFRA1
Molecular Function		
Cytokine binding	7.66×10^{-7}	IL2RB, CCR6, IL18BP, CMKLR1, CCR2, IL21R, CXCR6, CSF2RB, GFRA1
Nucleotide receptor activity	1.03×10^{-4}	P2RY12, P2RY13, P2RX7, P2RY10, GPR18
Purinergic nucleotide receptor activity	1.03×10^{-4}	P2RY12, P2RY13, P2RX7, P2RY10, GPR18
Cytokine activity	3.77×10^{-4}	TNF, CCL2, CCL19, FASLG, CXCL6, IL33, GREM1, LTB
Chemokine receptor activity	6.42×10^{-4}	CCR6, CMKLR1, CCR2, CXCR6

Table I. Continued.

B, Upregulated genes

Category/GO term	P-value	Genes
Biological process		
Viral process	1.98×10^{-4}	DDX11, BTRC, RBM15B, TSC2, BRD4, TPR, RCC1
RNA processing	2.68×10^{-3}	DHX9, DDX54, HNRNPDL, DHX30
Cell division	1.76×10^{-2}	ANAPC1, BRCC3, TPR, RCC1, SMC1A
Positive regulation of chromatin binding	2.54×10^{-2}	KDM1A, DDX11
DNA repair	2.95×10^{-2}	BRCC3, DDX11, TDP1, SMC1A
Cellular Component		
Nucleoplasm	2.85×10^{-5}	ANAPC1, DHX9, BRCC3, ZMYM3, TONSL, BTRC, RBM15B, SNIP1, HNRNPDL, RCC1, BMS1, GTSE1, GPS2, KDM1A, DDX11, GTF2IRD1, BRD4, TPR, PPP4C, SMC1A, HDAC8
Cytoplasm	7.46×10^{-5}	SHROOM3, BTRC, SNIP1, RCC1, TIPRL, ZIC2, DDX11, DLG3, BRD4, TPR, DHX30, PPP4C, SAMD4B, DHX9, BRCC3, ZMYM3, PIK3C2A, TONSL, AMBRA1, HNRNPDL, MID1, GCN1, GTF2IRD1, TDP1, TSC2, USP47, SPTBN1, SMC1A, HDAC8
Condensed nuclear chromosome	2.77×10^{-3}	BRD4, RCC1, SMC1A
Nucleus	2.86×10^{-3}	BTRC, SNIP1, RCC1, BMS1, ZIC2, KDM1A, DDX11, BRD4, DHX30, PPP4C, TPR, SAMD4B, DHX9, PGAP2, BRCC3, PIK3C2A, CIZ1, HNRNPDL, GTF2IRD1, TDP1, TSC2, USP47, SPTBN1, DDX54, SMC1A, HDAC8
Cytoplasmic microtubule	8.97×10^{-3}	SPACA9, MID1, GTSE1
Molecular Function		
Poly(A) RNA binding	3.79×10^{-4}	DHX9, RBM15B, SNIP1, SPTBN1, DDX54, HNRNPDL, TPR, DHX30, SMC1A, BMS1, GCN1, SAMD4B
Chromatin binding	9.98×10^{-4}	KDM1A, DDX11, BRD4, TPR, RCC1, DHX30, SMC1A
ATP-dependent helicase activity	3.32×10^{-3}	DHX9, DDX11, DHX30
Protein binding	7.47×10^{-3}	REPS1, BTRC, RBM15B, SNIP1, RCC1, TIPRL, GTSE1, KDM1A, DDX11, ILVBL, DLG3, PEX13, BRD4, TPR, DHX30, PPP4C, DHX9, PGAP2, CHDH, BRCC3, SPACA9, TONSL, CIZ1, AMBRA1, HNRNPDL, MID1, GPS2, ARHGAP32, TDP1, TSC2, USP47, SPTBN1, SMC1A, HDAC8, GATC, IQCE
Nucleic acid binding	7.89×10^{-3}	DHX9, DDX11, CIZ1, RBM15B, GPATCH1, DDX54, HNRNPDL, DHX30, ZIC2

GO, Gene Ontology.

Table II. KEGG pathway analysis of differentially expressed genes in the GSE18549 dataset.

KEGG pathway	P-value	Genes
Downregulated genes		
Cytokine-cytokine receptor interaction	3.90×10^{-7}	IL2RB, CCL2, TNF, IL21R, FASLG, TNFRSF17, CCL19, CXCL6, CCR6, CXCR6, CCR2, CSF2RB, LTB, CD27
Natural killer cell-mediated cytotoxicity	2.36×10^{-5}	CD48, ITGAL, CD244, SH2D1A, TNF, CD247, FASLG, PRKCB, LCP2
Hematopoietic cell lineage	1.08×10^{-3}	CR1, TNF, CR2, CD3G, CD34, MS4A1
Chemokine signaling pathway	6.98×10^{-3}	CCR6, CCL2, CCR2, CXCR6, CCL19, CXCL6, PRKCB
T-cell receptor signaling pathway	1.70×10^{-2}	TNF, CD3G, CD247, CD28, LCP2
Upregulated genes		
Oocyte meiosis	3.89×10^{-2}	ANAPC1, BTRC, SMC1A

KEGG, Kyoto Encyclopedia of Genes and Genomes.
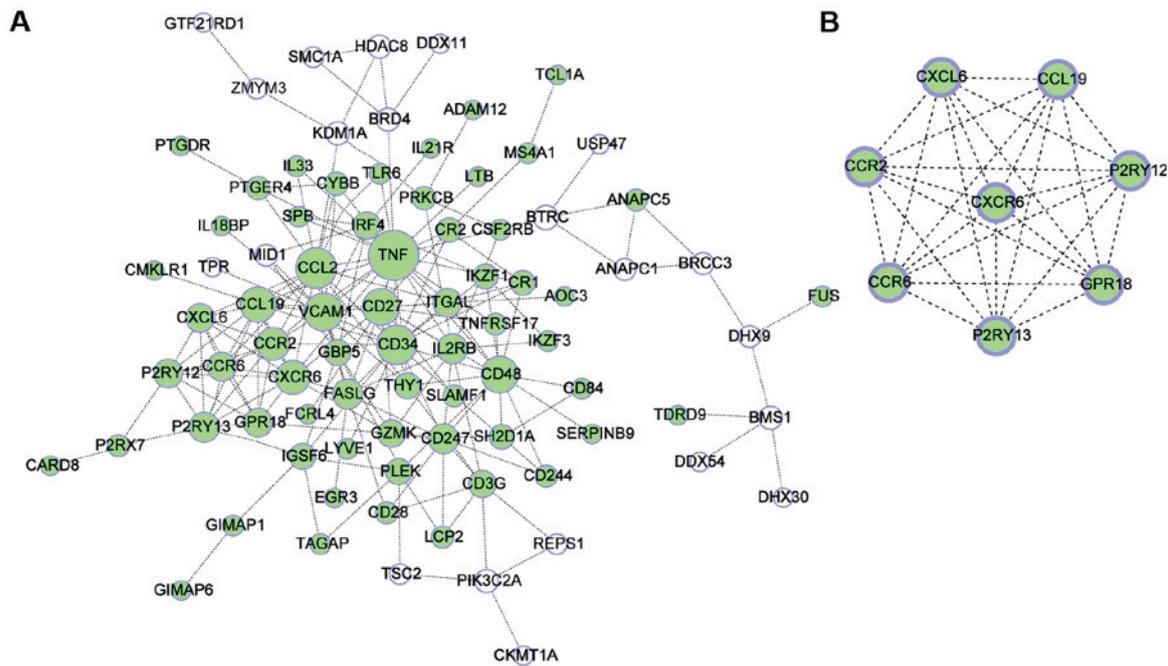
Figure 2. PPI network and a significant module constructed from the DEGs in dataset GSE18549. (A) Using the STRING online database, a total of 190 DEGs were filtered into the PPI network complex. (B) A significant module selected from the PPI network. The nodes represent proteins, the edges represent the interaction of proteins, the green circles represent downregulated genes and the white circles represent the upregulated genes. PPI, protein-protein interaction; DEG, differentially expressed gene.
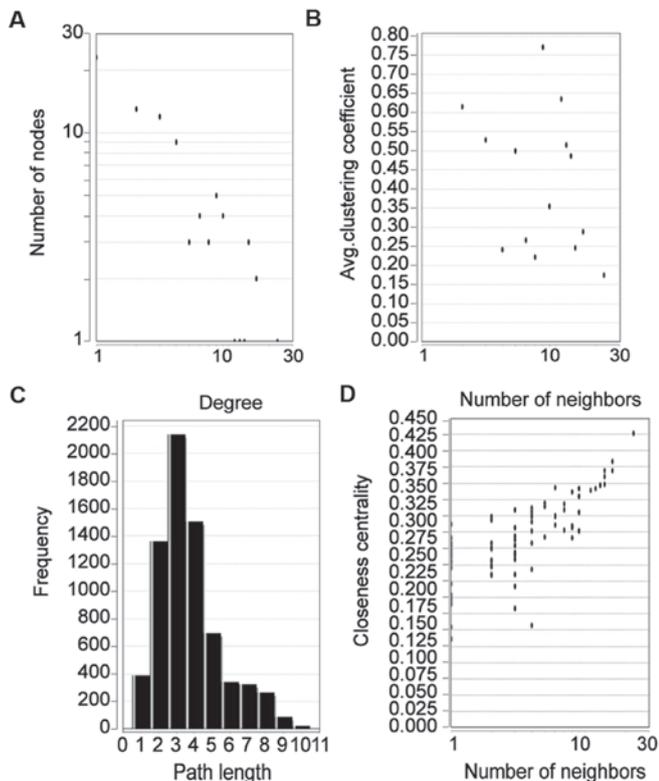
Figure 3. Topological properties of the protein-protein interaction network. (A) Node-degree distribution; (B) average clustering coefficient distribution; (C) shortest path length distribution; (D) closeness centrality.

metastasis and brain metastasis. A total of 190 DEGs were obtained, which included 129 downregulated genes and 61 upregulated genes. To further investigate the functions of the

DEGs, GO functional annotation and KEGG pathway enrichment analysis were used based on the DAVID database. The GO analysis demonstrated that downregulated DEGs were mainly associated with 'immune response', 'cell activation' and 'leukocyte activation', which was in accordance with previous studies. For instance, secondary brain cancers frequently exhibit high expression of programmed cell death 1 (PD-1) ligand 1, which may be inhibited by novel treatments that activate the immune system (24,25). Furthermore, degraded white-matter tract integrity in the areas with high T-cell densities are thought to provide active microenvironments for brain metastases (26); in addition, immune checkpoint inhibitors targeting PD-1 and cytotoxic T lymphocyte-associated protein 4 are becoming a frontline therapy in melanoma, which also suggests that the immune response is involved in the development of brain metastases (27). On the contrary, the upregulated DEGs were involved in 'DNA repair' and 'viral process', which were also consistent with previous data. Overexpression of DNA repair genes BRCA1-associated RING domain 1 and RAD51 recombinase frequently occur in brain metastases from breast cancer (28). The DNA-damage response pathway was also determined to be involved in leptomeningeal metastasis of non-small cell lung cancer (29). Furthermore, high rates of DNA repair mutations were identified in brain metastases from prostate cancer (30). In addition, the prevalence of human cytomegalovirus proteins and nucleic acids is high in primary and metastatic tumors, indicating that this virus may drive the development of metastatic brain tumors (31).

The results of the KEGG analysis indicated that the downregulated DEGs were mainly enriched in 'chemokine signaling pathway'. Previous studies demonstrated that inflammatory chemokines and their receptors regulate tumor cell migration and participate in tumor growth, metastasis, angiogenesis and

Table III. GO analysis of genes in the significant module.

Category/GO term	P-value	Genes
Biological Process		
G-protein coupled receptor protein signaling pathway	2.67×10^{-8}	P2RY12, P2RY13, CCR6, GPR18, CCR2, CXCR6, CCL19, CXCL6
Cell surface receptor linked signal transduction	9.06×10^{-7}	P2RY12, P2RY13, CCR6, GPR18, CCR2, CXCR6, CCL19, CXCL6
Taxis	5.49×10^{-5}	CCR6, CCR2, CCL19, CXCL6
Chemotaxis	5.49×10^{-5}	CCR6, CCR2, CCL19, CXCL6
Locomotory behavior	2.71×10^{-4}	CCR6, CCR2, CCL19, CXCL6
Cellular Component		
Plasma membrane	2.69×10^{-2}	P2RY12, P2RY13, CCR6, GPR18, CCR2, CXCR6
Molecular Function		
C-C chemokine receptor activity	2.98×10^{-5}	CCR6, CCR2, CXCR6
C-C chemokine binding	2.98×10^{-5}	CCR6, CCR2, CXCR6
Chemokine receptor activity	7.43×10^{-5}	CCR6, CCR2, CXCR6
Chemokine binding	8.69×10^{-5}	CCR6, CCR2, CXCR6
Nucleotide receptor activity, G-protein coupled	1.00×10^{-4}	P2RY12, P2RY13, GPR18

GO, Gene Ontology.

Table IV. KEGG pathway analysis of genes in the significant module.

KEGG pathway	Count	P-Value	Genes
hsa04062:Chemokine signaling pathway	5	8.61×10^{-6}	CCR6, CCR2, CXCR6, CCL19, CXCL6
hsa04060:Cytokine-cytokine receptor interaction	5	3.31×10^{-5}	CCR6, CCR2, CXCR6, CCL19, CXCL6

Hsa, *Homo sapiens*; KEGG, Kyoto Encyclopedia of Genes and Genomes; CXCL6, chemokine (C-X-C) ligand 6; CCL, C-C motif chemokine ligand; CCR, C-C chemokine receptor.

invasion through the interaction between mesenchymal cells and neoplastic cells (32,33). Of note, the upregulated DEGs were mainly associated with 'Oocyte meiosis'. However, there is no evidence to support that this pathway was associated with BMLC or brain metastasis from other types of tumor.

Furthermore, the protein interactions among the screened DEGs were predicted. In the PPI network, 9 hub genes with the highest degree of connectivity were selected, which included TNF, CCL2, CD34, VCAM1, CD48, CD27, CCL19, CXCR6 and CCR2. TNF encodes a multifunctional proinflammatory cytokine that belongs to the TNF superfamily. TNF-α has an important role in the adhesion of non-small cell lung cancer cells to brain endothelium mediated by CD62E (34). A previous study suggested that microRNA (miR)-509 has a critical role in brain metastasis of breast cancer by modulating the Ras homolog family member C/TNF-α network (35). Among the above genes, half of the hub genes, including CCL2, CCL19, CCR2 and CXCR6, were involved in the chemokine signaling pathway. CCL2, a small cytokine that belongs to the CC chemokine family, is anchored in the plasma membrane of endothelial cells by glycosaminoglycan side chains of proteoglycans; CCL2 exhibits a chemotactic activity for monocytes and basophils. miR-19a contained in astrocyte-derived exosomes reversibly downregulates phosphatase and tension homolog in tumor cells, resulting in secretion of CCL2, which is able to recruit brain metastasis-promoting myeloid cells (36). Conceivably, CCR2, as one of the receptors for CCL2, mediated the roles of CCL2 in breast cancer metastasis (37). Similar to CCL2, CCL19 is another member of the CC motif chemokine superfamily and functions as a tumor suppressor in lung cancer; CCL19-expressing fibroblastic stromal cells were indicated to inhibit lung carcinoma growth by promoting local anti-tumor T-cell responses (38). CCL19 exhibited anti-tumor effects by promoting interferon-γ-dependent responses in a lung cancer model (39); however, the roles of CCL19 in BMLC have remained to be determined. CXCR6, belonging to family A of the G protein-coupled receptor superfamily, was identified to be associated with B-lineage maturation and antigen-driven B-cell differentiation. As a binding partner, CXCR6 mediates CXCL16-associated signaling. Recently, several studies indicated that CXCL16/CXCR6 signaling drives metastasis of different cancer types by promoting a protumor inflammatory environment or attracting cancer cells (40,41).

Another group of hub genes, CD, are a defined subset of cellular surface receptors, which include CD48, CD27 and CD34. CD48, also known as B-lymphocyte activation marker or signaling lymphocytic activation molecule 2, encodes a member of the CD2 subfamily of immunoglobulin-like receptors. Activation of CD48 by injecting anti-CD48 monoclonal antibodies resulted in the inhibition of tumor metastasis of melanoma (42). CD27, a member of the TNF-receptor superfamily, is required for generation and long-term maintenance of T-cell immunity; Pagès *et al* (43) discovered that the expression of CD27 was associated with early metastasis in colorectal cancer. CD34, a single-pass membrane protein, is highly glycosylated and phosphorylated by protein kinase C and expressed on human hematopoietic progenitor cells and the small-vessel endothelium of a variety of tissue types (44); Upregulation of CD34-positive lymphatic/vascular endothelial progenitor cells are associated with metastasis of ovarian cancer (45).

VCAM1 is an Ig-like cell adhesion molecule expressed by cytokine-activated endothelium (46). One study based on RNA sequencing analysis confirmed that the expression of VCAM1 was upregulated in brain tissue harboring metastases, which suggested that this gene may contribute to the establishment of brain metastases from breast cancer or melanoma (47); however, the functions VCAM1 in BMLC have so far remained elusive.

In conclusion, although the present study had certain limitations, including the small number of cases and the lack of validation in clinical samples, the present analysis identified distinct key genes and pathways closely associated with BMLC, which may contribute to the current knowledge of the complex mechanisms of BMLC. Of note, the present results warrant confirmation by further investigations.

Availability of data and materials

The datasets used and/or analyzed in the present study are available from the corresponding authors on reasonable request.

Authors' contributions

FR and PW designed the experiments, and XZ and NW collected and analyzed the data. TC and HG wrote the manuscript. YL and HC downloaded the gene expression profile from the GEO. All authors critically reviewed the content and approved the final version for publication.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Siegel RL, Miller KD and Jemal A: Cancer statistics, 2018. CA Cancer J Clin 68: 7-30, 2018.
2. Chuang CH, Greenside PG, Rogers ZN, Brady JJ, Yang D, Ma RK, Caswell DR, Chiou SH, Winters AF, Grüner BM, *et al*: Molecular definition of a metastatic lung cancer state reveals a targetable CD109-Janus kinase-Stat axis. Nat Med 23: 291-300, 2017.
3. Tamura T, Kurishima K, Nakazawa K, Kagohashi K, Ishikawa H and Hizawa N: Specific organ metastasis and survival in metastatic non-small-cell lung cancer. Mol Clin Oncol 3: 217-221, 2015.
4. Yano T, Yokoyama H, Inoue T, Asoh H, Tayama K, Takai E and Ichinose Y: The first site of recurrence after complete resection in non-small-cell carcinoma of the lung. Comparison between pN0 disease and pN2 disease. J Thorac Cardiovasc Surg 108: 680-683, 1994.
5. Chen W, Zhang S and Zou X: Estimation and projection of lung cancer incidence and mortality in China. Zhongguo Fei Ai Za Zhi 13: 488-493, 2010 (In Chinese).
6. Sandler A, Hirsh V, Reck M, von Pawel J, Akerley W and Johnson DH: An evidence-based review of the incidence of CNS bleeding with anti-VEGF therapy in non-small cell lung cancer patients with brain metastases. Lung Cancer 78: 1-7, 2012.
7. Linardou H, Kalogeras KT, Kronenwett R, Kouvatseas G, Wirtz RM, Zagouri F, Gogas H, Christodoulou C, Koutras AK, Samantas E, *et al*: The prognostic and predictive value of mRNA expression of vascular endothelial growth factor family members in breast cancer: A study in primary tumors of high-risk early breast cancer patients participating in a randomized Hellenic Cooperative Oncology Group trial. Breast Cancer Res 14: R145, 2012.
8. Chen G, Liu XY, Wang Z and Liu FY: Vascular endothelial growth factor C: The predicator of early recurrence in patients with N2 non-small-cell lung cancer. Eur J Cardiothorac Surg 37: 546-551, 2010.
9. Chen G, Wang Z, Liu XY and Liu FY: High-level CXCR4 expression correlates with brain-specific metastasis of non-small cell lung cancer. World J Surg 35: 56-61, 2011.
10. Han CH and Brastianos PK: Genetic characterization of brain metastases in the era of targeted therapy. Front Oncol 7: 230, 2017.
11. Tarca AL, Romero R and Draghici S: Analysis of microarray experiments of gene expression profiling. Am J Obstet Gynecol 195: 373-388, 2006.
12. Wong SY and Hynes RO: Lymphatic or hematogenous dissemination: How does a metastatic tumor cell decide? Cell Cycle 5: 812-817, 2006.
13. Naxerova K, Reiter JG, Brachtel E, Lennerz JK, van de Wetering M, Rowan A, Cai T, Clevers H, Swanton C, Nowak MA, *et al*: Origins of lymphatic and distant metastases in human colorectal cancer. Science 357: 55-60, 2017.
14. Law CW, Alhamdoosh M, Su S, Smyth GK and Ritchie ME: RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. F1000Res 5: 1408, 2016.
15. Gene Ontology Consortium: The gene ontology (GO) project in 2006. Nucleic Acids Res 34: D322-D326, 2006.

16. Kanehisa M, Furumichi M, Tanabe M, Sato Y and Morishima K: KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 45: D353-D361, 2017.
17. Huang da W, Sherman BT and Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4: 44-57, 2009.
18. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, et al: STRING v10: Protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43: D447-D452, 2015.
19. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T: Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res 13: 2498-2504, 2003.
20. Strogatz SH: Exploring complex networks. Nature 410: 268-276, 2001.
21. Mori T, Ikeda DD, Fukushima T, Takenoshita S and Kochi H: NIRF constitutes a nodal point in the cell cycle network and is a candidate tumor suppressor. Cell Cycle 10: 3284-3299, 2011.
22. Preusser M, Winkler F, Valiente M, Manegold C, Moyal E, Widhalm G, Tonn JC and Zielinski C: Recent advances in the biology and treatment of brain metastases of non-small cell lung cancer: Summary of a multidisciplinary roundtable discussion. ESMO Open 3: e000262, 2018.
23. Whitsett TG, Inge LJ, Dhruv HD, Cheung PY, Weiss GJ, Bremner RM, Winkles JA and Tran NL: Molecular determinants of lung cancer metastasis to the central nervous system. Transl Lung Cancer Res 2: 273-283, 2013.
24. Duchnowska R, Pęksa R, Radecka B, Mandat T, Trojanowski T, Jarosz B, Czartoryska-Arłukowicz B, Olszewski WP, Och W, Kalinka-Warzocha E, Kozłowski W, et al: Immune response in breast cancer brain metastases and their microenvironment: The role of the PD-1/PD-L axis. Breast Cancer Res 18: 43, 2016.
25. Berghoff AS, Venur VA, Preusser M and Ahluwalia MS: Immune checkpoint inhibitors in brain metastases: From biology to treatment. Am Soc Clin Oncol Educ Book 35: e116-e122, 2016.
26. Zakaria R, Platt-Higgins A, Rathi N, Radon M, Das S, Das K, Bhojak M, Brodbelt A, Chavredakis E, Jenkinson MD and Rudland PS: T-cell densities in brain metastases are associated with patient survival times and diffusion tensor MRI changes. Cancer Res 78: 610-616, 2018.
27. Taggart D, Andreou T, Scott KJ, Williams J, Rippaus N, Brownlie RJ, Ilett EJ, Salmond RJ, Melcher A and Lorger M: Anti-PD-1/anti-CTLA-4 efficacy in melanoma brain metastases depends on extracranial disease and augmentation of CD8+ T cell trafficking. Proc Natl Acad Sci USA 115: E1540-E1549, 2018.
28. Woditschka S, Evans L, Duchnowska R, Reed LT, Palmier D, Qian Y, Badve S, Sledge G Jr, Gril B, Aladjem MI, et al: DNA double-strand break repair genes and oxidative damage in brain metastasis of breast cancer. J Natl Cancer Inst 106: diu145, 2014.
29. Fan Y, Zhu X, Xu Y, Lu X, Xu Y, Wang M, Xu H, Ding J, Ye X, Fang L, et al: Cell-cycle and DNA-damage response pathway is involved in leptomeningeal metastasis of non-small cell lung cancer. Clin Cancer Res 24: 209-216, 2018.
30. Pritchard CC, Mateo J, Walsh MF, De Sarkar N, Abida W, Beltran H, Garofalo A, Gulati R, Carreira S, Eeles R, et al: Inherited DNA-repair gene mutations in men with metastatic prostate cancer. N Engl J Med 375: 443-453, 2016.
31. Taher C, Frisk G, Fuentes S, Religa P, Costa H, Assinger A, Vetvik KK, Bukholm IR, Yaiw KC, Smedby KE, et al: High prevalence of human cytomegalovirus in brain metastases of patients with primary breast and colorectal cancers. Transl Oncol 7: 732-740, 2014.
32. Cheng ZH, Shi YX, Yuan M, Xiong D, Zheng JH and Zhang ZY: Chemokines and their receptors in lung cancer progression and metastasis. J Zhejiang Univ Sci B 17: 342-351, 2016.
33. Reckamp KL, Strieter RM and Figlin RA: Chemokines as therapeutic targets in renal cell carcinoma. Expert Rev Anticancer Ther 8: 887-893, 2008.
34. Jassam SA, Maherally Z, Smith JR, Ashkg K, Roncaroli F, Fillmore HL and Pilkington GJ: TNF-α enhancement of CD62E mediates adhesion of non-small cell lung cancer cells to brain endothelium via CD15 in lung-brain metastasis. Neuro Oncol 18: 679-690, 2016.
35. Xing F, Sharma S, Liu Y, Mo YY, Wu K, Zhang YY, Pochampally R, Martinez LA, Lo HW and Watabe K: miR-509 suppresses brain metastasis of breast cancer cells by modulating RhoC and TNF-α. Oncogene 34: 4890-4900, 2015.
36. Zhang L, Zhang S, Yao J, Lowery FJ, Zhang Q, Huang WC, Li P, Li M, Wang X, Zhang C, et al: Microenvironment-induced PTEN loss by exosomal microRNA primes brain metastasis outgrowth. Nature 527: 100-104, 2015.
37. Lu X and Kang Y: Chemokine (C-C motif) ligand 2 engages CCR2+ stromal cells of monocytic origin to promote breast cancer metastasis to lung and bone. J Biol Chem 284: 29087-29096, 2009.
38. Cheng HW, Onder L, Cupovic J, Boesch M, Novkovic M, Pikor N, Tarantino I, Rodriguez R, Schneider T, Jochum W, et al: CCL19-producing fibroblastic stromal cells restrain lung carcinoma growth by promoting local antitumor T-cell responses. J Allergy Clin Immunol 142: 1257-1271, 2018.
39. Hillinger S, Yang SC, Zhu L, Huang M, Duckett R, Atianzar K, Batra RK, Strieter RM, Dubinett SM and Sharma S: EBV-induced molecule 1 ligand chemokine (ELC/CCL19) promotes IFN-gamma-dependent antitumor responses in a lung cancer model. J Immunol 171: 6457-6465, 2003.
40. Chung B, Esmaeili AA, Gopalakrishna-Pillai S, Murad JP, Andersen ES, Kumar Reddy N, Srinivasan G, Armstrong B, Chu C, Kim Y, et al: Human brain metastatic stroma attracts breast cancer cells via chemokines CXCL16 and CXCL12. NPJ Breast Cancer 3: 6, 2017.
41. Gao Q, Zhao YJ, Wang XY, Qiu SJ, Shi YH, Sun J, Yi Y, Sun JY, Shi GM, Ding ZB, et al: CXCR6 upregulation contributes to a proinflammatory tumor microenvironment that drives metastasis and poor patient outcomes in hepatocellular carcinoma. Cancer Res 72: 3546-3556, 2012.
42. Johnson LA, Vaidya SV, Goldfarb RH and Mathew PA: 2B4 (CD244)-mediated activation of NK cells reduces metastases of B16F10 melanoma in mice. Anticancer Res 23: 3651-3655, 2003.
43. Pagès F, Berger A, Camus M, Sanchez-Cabo F, Costes A, Molidor R, Mlecnik B, Kirilovsky A, Nilsson M, Damotte D, et al: Effector memory T cells, early metastasis, and survival in colorectal cancer. N Engl J Med 353: 2654-2666, 2005.
44. Simmons DL, Satterthwaite AB, Tenen DG and Seed B: Molecular cloning of a cDNA encoding CD34, a sialomucin of human hematopoietic stem cells. J Immunol 148: 267-271, 1992.
45. Qiu H, Cao L, Wang D, Xu H and Liang Z: High levels of circulating CD34+/VEGFR3+ lymphatic/vascular endothelial progenitor cells is correlated with lymph node metastasis in patients with epithelial ovarian cancer. J Obstet Gynaecol Res 39: 1268-1275, 2013.
46. Lawson C, Ainsworth M, Yacoub M and Rose M: Ligation of ICAM-1 on endothelial cells leads to expression of VCAM-1 via a nuclear factor-kappaB-independent mechanism. J Immunol 162: 2990-2996, 1999.
47. Sato R, Nakano T, Hosonaga M, Sampetrean O, Harigai R, Sasaki T, Koya I, Okano H, Kudoh J, Saya H and Arima Y: RNA sequencing analysis reveals interactions between breast cancer or melanoma cells and the tissue microenvironment during brain metastasis. Biomed Res Int 2017: 8032910, 2017.