Microarray based analysis of gene regulation by mesenchymal stem cells in breast cancer

MING ZHANG1*, CHANG E. GAO2*, WEN HUI LI1, YI YANG1, LI CHANG1, JIAN DONG3, YAN XIN REN4 and DE DIAN CHEN5

1Department of Radiation Oncology, The Third Affiliated Hospital of Kunming Medical University, Kunming, Yunnan 650118; 2Department of Medical Oncology, The First Affiliated Hospital of Kunming Medical University, Kunming, Yunnan 650032; 3Institute of Oncology, The Third Affiliated Hospital of Kunming Medical University; Departments of 4Head and Neck Surgery, and 5Breast Surgery, The Third Affiliated Hospital of Kunming Medical University, Kunming, Yunnan 650118, P.R. China

Received July 8, 2015; Accepted November 25, 2016

DOI: 10.3892/ol.2017.5776

Abstract. Breast cancer is one of the most common malignant tumors with a high case-fatality rate among women. The present study aimed to investigate the effects of mesenchymal stem cells (MSCs) on breast cancer by exploring the potential underlying molecular mechanisms. The expression profile of GSE43306, which refers to MDA-MB-231 cells with or without a 1:1 ratio of MSCs, was downloaded from Gene Expression Omnibus database for differentially expressed gene (DEG) screening. The Database for Annotation, Visualization and Integrated Discovery was used for gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis for DEGs. The protein-protein interactional (PPI) network of DEGs was constructed using the Search Tool for the Retrieval of Interacting Genes/Proteins. The data was subsequently analyzed using molecular complex detection for sub-network mining of modules. Finally, DEGs in modules were analyzed using GO and KEGG pathway enrichment analyses. A total of 291 DEGs including 193 upregulated and 98 downregulated DEGs were obtained. Uregulated DEGs were primarily enriched in pathways including response to wounding (P=5.92x10^-4), inflammatory response (P=5.92x10^-4) and defense response (P=1.20x10^-2), whereas downregulated DEGs were enriched in pathways including the cell cycle (P=7.13x10^-4), mitotic cell cycle (P=6.81x10^-2) and M phase (P=1.72 x10^-2). The PPI network, which contained 156 nodes and 289 edges, was constructed, and Fos was the hub node with the degree of 29. A total of 3 modules were mined from the PPI network. In total, 14 DEGs in module A were primarily enriched in GO terms, including response to wounding (P=4.77x10^-4), wounding healing (P=6.25x10^-7) and coagulation (P=1.13 x10^-7), and these DEGs were also enriched in 1 KEGG pathway (complement and coagulation cascades; P=0.0036). Therefore, MSCs were demonstrated to exhibit potentially beneficial effects for breast cancer therapy. In addition, the screened DEGs, particularly in PPI network modules, including FN1, CD44, NGF, SERPINE1 and CCNA2, may be the potential target genes of MSC therapy for breast cancer.

Introduction

Breast cancer is one of the most common malignant tumors and has a high case-mortality rate among women, accounting for ~1/3 cancer cases diagnosed in the United States (1). Breast cancer has a heredity element and frequently occurs during menopause (2). Signs of breast cancer include changes in breast shape and skin, a breast lump and cyst fluid discharged from the nipple (3). At present, the treatments for breast cancer primarily include tumor resection, radiation treatment and chemotherapy (4,5). However, these treatment methods are associated with a high risk of recurrence (6). Therefore, the identification, development and study of novel treatment methods is required.

Mesenchymal stem cells (MSCs) are an important type of adult stem cells, which serve a role in the processes of tumor growth and metastasis (7). In vitro, MSCs are able to arrest the cell cycle progression of tumor cells in the G1 phase and reduce their apoptotic rate (8). In addition, previous studies have revealed that MSCs possess a number of functions, including hematopoietic support, immunoregulation, multilineage differentiation and specific migration (9,10). Studeny et al (11) demonstrated that MSCs are able to form an effective platform for the local production of interferon (IFN)-β, suppressing the process of pulmonary metastasis. Furthermore, MSCs may be used as carriers of a number of therapeutics, including interleukin (IL)-2, IFNs and C-X3-C motif chemokine receptor 1, in order to induce apoptosis and inhibit tumor cell differentiation (12-14).
A previous study revealed that MSCs are able integrate into the tumor-associated stroma and to affect the development of breast cancer (15). In co-culture with breast cancer cells, MSCs secreted a series of factors, including chemokine C-C motif chemokine ligand 2, IL-6 and tissue inhibitor of metalloproteinase 1 (TIMP-1), restricting the growth of cancer cells (16,17). Furthermore, MSCs were able to inhibit the proliferation of breast cancer cells by secreting Dickkopf-related protein 1, a novel inhibitor of Wnt signaling (18). In addition, MSCs are able to arrest cells at the G0/G1 phase of the cell cycle through upregulation of tumor protein 21 (p21) and caspase-3, further inhibiting cell growth (19). Furthermore, breast cancer-associated lymphedema of the arm can be effectively treated using autologous bone MSC transplantation (20).

However, certain opposing studies have revealed that MSCs are able to promote breast cancer metastasis (21,22). Therefore, the objective of the current study was to further examine this controversial issue by analyzing the underlying molecular mechanisms of the effect of MSCs in breast cancer. In the present study, to further explore the molecular mechanisms of MSCs in breast cancer, the GSE43306 gene expression dataset was downloaded from the Gene Expression Omnibus (GEO) database for differentially expressed gene (DEG) screening, pathway enrichment analysis and protein-protein interaction (PPI) network construction. Finally, PPI network modules were screened and analyzed.

Materials and methods

Microarray data. The GSE43306 expression profile dataset was downloaded from the GEO database (http://www.ncbi.nlm.nih.gov/geo/). This dataset was collected using Illumina HiSeq 2000 (23). A total of nine tissue samples consisting of MDA-MB-231 cells comprised this dataset, including five samples supplemented with a 1:1 ratio of bone marrow MSCs (M) to MDA-MB-231 cells and four samples without MSC supplementation (A), which were compared in the current study to elucidate the effect of MSCs in breast cancer.

Data pre-processing and DEG analysis. The raw data were converted into a FASTQ format, subsequently the Next Generation Sequencing Quality Control Toolkit (24) was used for quality control and filtering of high-quality reads. Reads with ≥20 bases (70% read length) were selected as high-quality reads. By default, the high-quality reads were aligned to the full Human Genome (version 19) (https://www.ncbi.nlm.nih.gov/genome/51) using TopHat2 (25). Based on the results of comparison and genome annotation profiles, gene expression levels were calculated and genes with an expression value of 0 were removed. Subsequently, a gene expression matrix was produced through the removal of repeated gene symbols. The NOISeq R/Bioc package (bioconductor.org) (26) was applied to screen for DEGs between the M and A groups with the threshold of q<0.05.

Functional and pathway enrichment analysis. The Database for Annotation, Visualization and Integrated Discovery, which uses analytical tools to extract biological functions for numerous genes (27), was used for Gene Ontology (GO) (28) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (29) functional and pathway enrichment analysis, respectively, of the DEGs identified. The cut-off criteria were determined as P<0.05 and an enriched gene count ≥2.

PPI network construction. A PPI network of the DEGs was constructed using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database, with a threshold combined score of >0.4. In the PPI network, the nodes and edges (lines) represent proteins and their interactions, respectively. Nodes with an average node connective degree ≥9 were regarded as hub proteins (30).

PPI network analysis. The PPI network was analyzed using Molecular Complex Detection (31), which mines for sub-network modules. The sub-networks were screened with the following default cut-off thresholds: Degree, 2; node score, 0.2; K-core, 2; maximum depth, 100. Subsequently, the nodes with the average node degree ≥3 and combined score (consisting of neighborhood, fusion, co-occurrence, co-expression, experimental database and text mining) were screened, and DEGs in these modules were used for pathway enrichment analysis.

Results

Screening of differentially expressed genes. Following pre-processing of the GSE43306 dataset, a total of 291 DEGs, including 193 upregulated and 98 downregulated DEGs, were identified (data not shown).

Functional and pathway enrichment analysis. As presented in Table 1, upregulated DEGs were primarily enriched in the following two types of GO terms: The first type included wound response (P=5.92x10^-7), the inflammatory response (P=5.92x10^-7) and the immune response (P=1.20x10^-2); and the second type included wound healing (P=4.97x10^-6), regulation of body fluid levels (P=6.82x10^-6), coagulation (P=6.83x10^-4), blood coagulation (P=6.83x10^-4) and hemostasis (P=9.23x10^-4). Simultaneously, the upregulated DEGs were enriched in a number of other functions, including the mitogen-activated protein kinase signaling pathway (P=4.56x10^-5) and hematopoietic cell lineage (P=2.13x10^-7). The downregulated DEGs were also enriched in two types of GO term, as follows: The first type included the cell cycle (P=7.13x10^-4), mitotic cell cycle (P=6.81x10^-4), M phase (P=1.72x10^-3), M phase of the mitotic cell cycle (P=1.93x10^-4), cell cycle phase (P=4.08x10^-5) and cell cycle process (P=4.41x10^-2); the second type included RNA processing (P=3.35x10^-5), RNA splicing (P=9.61x10^-3), mRNA processing (P=1.57x10^-5) and mRNA metabolic process (P=2.70x10^-5).

PPI network. The PPI network constructed contained 156 proteins and 289 interactions (Fig. 1). Based on the average degree of the nodes, the following 12 nodes with degrees ≥9 were obtained: Fos proto-oncogene AP-1 transcription factor subunit (FOS), vascular endothelial growth factor A (VEGFA), fibronectin 1 (FN1), cluster of differentiation 44 (CD44), nerve growth factor (NGF), activating transcription factor 3 Serpin Family E Member 1 (SERPINE1), cyclin A2 (CCNA2), PBZ binding kinase, tissue factor F3, heme
oxygenase-1 and ferritin heavy chain 1. Among them, FOS and VEGFA were the hub proteins with the highest node degree (29).

Analysis of PPI network modules. Through module analysis of the PPI network, a total of 3 modules, including modules A, B and C, were obtained (data not shown). The number of nodes in modules A, B and C were 14, 5 and 6, respectively. While the number of interactions in modules A, B and C were 33, 8 and 9, respectively. Module A is illustrated in Fig. 2. Amongst the three modules, module A had the highest enriched score. Therefore, module A was further analyzed with GO functional and KEGG pathway enrichment analyses, respectively (Table II). The 14 DEGs in module A were primarily enriched in GO terms such as wound response (P=4.77x10^-6), wound healing (P=6.25x10^-6) and coagulation (P=1.13x10^-7). In addition, these DEGs were enriched in one KEGG pathway, the complement and coagulation cascade (P=0.0036).

Table I. Top two types of functional and pathway enrichment analysis for upregulated and downregulated DEGs through GO and KEGG.

<table>
<thead>
<tr>
<th>Category</th>
<th>GO or KEGG term</th>
<th>Description</th>
<th>Number of nodes</th>
<th>P-value</th>
<th>Enrichment score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up 1</td>
<td>BP GO:0009611</td>
<td>Wound response</td>
<td>21</td>
<td>5.918x10^-7</td>
<td>3.792</td>
</tr>
<tr>
<td>Up 1</td>
<td>BP GO:0006954</td>
<td>Inflammatory response</td>
<td>12</td>
<td>5.948x10^-4</td>
<td></td>
</tr>
<tr>
<td>Up 1</td>
<td>BP GO:0006952</td>
<td>Defense response</td>
<td>14</td>
<td>1.199x10^-2</td>
<td></td>
</tr>
<tr>
<td>Up 2</td>
<td>BP GO:0042060</td>
<td>Wound healing</td>
<td>12</td>
<td>4.970x10^-6</td>
<td>3.567</td>
</tr>
<tr>
<td>Up 2</td>
<td>BP GO:0050878</td>
<td>Regulation of body fluid levels</td>
<td>8</td>
<td>6.818x10^-4</td>
<td></td>
</tr>
<tr>
<td>Up 2</td>
<td>BP GO:0050817</td>
<td>Coagulation</td>
<td>7</td>
<td>6.834x10^-4</td>
<td></td>
</tr>
<tr>
<td>Up 2</td>
<td>BP GO:0007596</td>
<td>Blood coagulation</td>
<td>7</td>
<td>6.834x10^-4</td>
<td></td>
</tr>
<tr>
<td>Up 2</td>
<td>BP GO:0007599</td>
<td>Hemostasis</td>
<td>7</td>
<td>9.229x10^-4</td>
<td></td>
</tr>
<tr>
<td>Down 1</td>
<td>BP GO:00007049</td>
<td>Cell cycle</td>
<td>12</td>
<td>7.129x10^-4</td>
<td>1.758</td>
</tr>
<tr>
<td>Down 1</td>
<td>BP GO:0000278</td>
<td>Mitotic cell cycle</td>
<td>7</td>
<td>6.088x10^-3</td>
<td></td>
</tr>
<tr>
<td>Down 1</td>
<td>BP GO:0000279</td>
<td>M phase</td>
<td>6</td>
<td>1.724x10^-2</td>
<td></td>
</tr>
<tr>
<td>Down 1</td>
<td>BP GO:0000087</td>
<td>M phase of mitotic cell cycle</td>
<td>5</td>
<td>1.932x10^-2</td>
<td></td>
</tr>
<tr>
<td>Down 1</td>
<td>BP GO:0022403</td>
<td>Cell cycle phase</td>
<td>6</td>
<td>4.085x10^-2</td>
<td></td>
</tr>
<tr>
<td>Down 1</td>
<td>BP GO:0022402</td>
<td>Cell cycle process</td>
<td>7</td>
<td>4.415x10^-2</td>
<td></td>
</tr>
<tr>
<td>Down 2</td>
<td>BP GO:0006396</td>
<td>RNA processing</td>
<td>9</td>
<td>3.555x10^-3</td>
<td>1.758</td>
</tr>
<tr>
<td>Down 2</td>
<td>BP GO:0006397</td>
<td>mRNA processing</td>
<td>6</td>
<td>9.611x10^-3</td>
<td></td>
</tr>
<tr>
<td>Down 2</td>
<td>BP GO:0016071</td>
<td>mRNA metabolic process</td>
<td>6</td>
<td>1.566x10^-2</td>
<td></td>
</tr>
<tr>
<td>Down 2</td>
<td>BP GO:0000375</td>
<td>RNA splicing, via transesterification reactions</td>
<td>4</td>
<td>2.702x10^-2</td>
<td></td>
</tr>
<tr>
<td>Down 2</td>
<td>BP GO:0000377</td>
<td>RNA splicing, via transesterification reactions with bulged adenosine as a nucleophile</td>
<td>4</td>
<td>3.303x10^-2</td>
<td></td>
</tr>
<tr>
<td>Down 2</td>
<td>BP GO:0000398</td>
<td>Nuclear mRNA splicing, via the spliceosome</td>
<td>4</td>
<td>3.303x10^-2</td>
<td></td>
</tr>
</tbody>
</table>

Up, upregulated; down, downregulated; BP, biological process; P, pathway; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; MAPK, mitogen-activated protein kinase; DEG, differentially expressed gene.

Discussion
The effect of MSCs on the development and progression of tumors, including breast cancer, remains a subject of debate (19). It is important to study the effects of MSCs on tumor growth in order to develop novel therapies for the treatment of cancer (32). In the current study, a PPI network was constructed from DEGs in breast cancer. An important module with 14 nodes, including FNI, CD44, NGF, SERPINE1 and CCNA2, was identified through mining for modules in this network.

FNI encodes the protein fibronectin 1, which is involved in cell adhesion and migration processes, including wound healing, embryogenesis, blood coagulation, host pathogenic defense and metastasis (33). FNI is upregulated during the chondrogenic differentiation of MSCs and in various metastatic chondrosarcomas (34). In addition, FNI expression has been demonstrated to be closely associated with various migration processes, including wound healing, embryogenesis
and the metastasis of cancer cells (35). In addition, in the PC3 prostate cancer cell line, FN1 was able to interact with the cell membrane reporter(s), inhibiting focus formation and tumorigenesis (36). Similarly, the mRNA expression of FN1 in renal cancer cells is significantly increased compared with that in normal renal tissue (37). In the current study, FN1 was enriched in mRNA metabolism signaling pathways, which indicated that FN1 may have an important inhibitive role in the development of breast cancer via increased mRNA or protein expression levels.

CD44, which encodes a receptor for hyaluronic acid, participates in various cellular functions including hematopoiesis, lymphocyte activation and tumor cell metastasis (38). Furthermore, CD44, which is widely used as a MSC marker, has been demonstrated to serve a role in the migration of MSCs via migration assays and small interfering RNA experiments (39).
In addition, during the development and progression of breast cancer, hyaluronic acid-CD44 signaling was able to inhibit breast cancer cell metastasis through epithelial-stromal interactions (40). The results of the current study are in agreement with these previous studies, demonstrating that CD44 was upregulated in cancer tissue samples supplemented with MDA-MB-231 cells and MSCs. Furthermore, the results of the current study suggest that CD44 serves a role in the metastasis of breast cancer, as CD44 was observed to be enriched in the pathway of hemostasis. In addition, these results indicate that CD44 is a potential therapeutic target for the treatment of breast cancer with MSCs.

NGF is a member of the NGF-β family that encodes a secreted protein with nerve growth stimulating activity (41). In a previous study of glioma, MSCs were able to produce NGF, thus having an antitumor effect (42). In addition, NGF has been demonstrated to be an important regulator of breast cancer progression, inhibiting the progression of breast cancer through interactions with the p75 neurotrophin receptor and p140 tropomyosin receptor kinase A (43). In the current study, NGF was upregulated in samples supplemented with MSCs, suggesting that it is a potential target gene for the treatment of breast cancer with MSCs.

SERPINE1, a member of the serine protease inhibitor superfamily, encodes a protein that inhibits urokinase-type plasminogen activator (uPA) (44). In breast cancer, tumor severity may be associated with polymorphism of the plasminogen activator inhibitor type 1 4G/5G gene (45). In addition, uPA has been revealed to be associated with poor patient prognosis and tumor metastasis in breast cancer via various signaling pathways, including extracellular matrix breakdown (46). In addition, uPA may stimulate MSC migration via the ERK signaling pathway (47). These results indicate that uPA is a potential therapeutic target for MSC-mediated breast cancer treatment.

CCNA2 is a member of the highly conserved cyclin family, which binds to and activates cyclin-dependent kinases (CDKs) 1 and 2 in order to promote G1/S and G2/M cell cycle progression (48). Lysine-specific demethylase 1 promotes the development and aggressiveness of breast cancer through regulating CCNA2 expression levels (49). In addition, overexpression of CCNA2 was demonstrated to be associated with the poor prognosis of patients with breast cancer (50). However, few studies have investigated the expression of CCNA2 in breast cancer cell samples supplemented with MSCs. In the current study, CCNA2 was identified to be a downregulated in breast cancer cell samples supplemented with MSCs compared with cells without supplementation. Furthermore, CCNA2 was demonstrated to be enriched in the cell cycle signaling pathway and was a node in the PPI network constructed for the DEGs. These results suggest that CCNA2 may be a potential target gene for the treatment of breast cancer with MSCs.

In conclusion, the results of the present study indicate that MSCs have beneficial effects for the treatment breast cancer. The DEGs identified, particularly those in PPI network modules, including FN1, CD44, NGF, SERPINE1 and CCNA2, may be the potential target genes for the treatment of breast cancer with MSCs. However, these results require confirmation through further study.

Acknowledgements

The present study was supported by the Technology and Applied Basic Research Program of Yunnan province (grant no. 2012FB165), the Health and Technology program of Yunnan province (grant no. 2016NS010) and the Yunnan provincial doctoral graduate student award.

References


