High throughput analysis to identify key gene molecules that inhibit adipogenic differentiation and promote osteogenic differentiation of human mesenchymal stem cells

WU-XUN PENG[1], CHANG-HONG GAO[2] and GUO-BAO HUANG[3]

[1]Department of Orthopedics, The Affiliated Hospital of Guizhou Medical University, Guiyang, Guizhou 550004;
Departments of [2]Orthopedics and [3]Burn and Plastic Surgery, Jinan Central Hospital Affiliated to
Shandong University, Jinan, Shandong 250013, P.R. China

Abstract. The present study investigated the key genes, which cause switch from adipogenic to osteogenic differentiation of human mesenchymal stem cells (hMSCs). The transcriptomic profile of hMSCs samples were collected from Array Express database. Differential expression network was constructed by calculating the Pearson 's correlation coefficient and ranked according to their topological features. The top 5% genes with degree ≥2 were selected as ego genes. Following the KEGG pathway enrichment analysis and the relevant miRNAs prediction, the miRNA-mRNA-pathway networks were constructed by combining the miRNA-mRNA pairs and mRNA-pathway pairs together. In total, we obtained 84, 119, 94 and 97 ego-genes in B, BI, BT and BTI groups, and *DLGAP5*, *DLGAP5*, *NUSAP1* and *NDC80* were the ego-genes with the highest z-score of each group, respectively. Beginning from each ego-gene, we identified 2 significant ego-modules with gene size ≥4 in group BI, and the ego-genes were *PBK* and *NCOA3*, respectively. Through KEGG pathway analysis, we found that most of the pathways enriched by ego-genes were associated with gene replication and repair, and cell proliferation. According to the miRNA prediction results, we found that some of the predicted miRNAs have been validated to be the regulatory miRNAs of these corresponding mRNAs. Finally we constructed a miRNA-mRNA-pathway network by integrating the miRNA-mRNA and mRNA-pathway pairs together. The constructed network gives us a more comprehensive understanding of the mechanism of osteogenic differentiation of hMSCs.

Correspondence to: Dr Guo-Bao Huang, Department of Burn and Plastic Surgery, Jinan Central Hospital Affiliated to Shandong University, 105 Jiefang Road, Jinan, Shandong 250013, P.R. China
E-mail: huangguobao1816@126.com

Introduction

Osteoporosis is a kind of systemic bone disease that is characteristic of low bone mass and bone tissue microstructure abnormality, which leads to a successive enhancement in bone fragility and fracture risk (1). Human mesenchymal stem cells (hMSCs) possess the potential to differentiate into numerous cell lineages, such as adipocytes, osteoblasts, fibroblasts and chondrocytes. In bone marrow of osteoporosis patients, the number of adipocytes are often increased, concomitant with a decrease in the pool of hMSC which could differentiate into osteoblasts (2,3). It has been confirmed that transforming growth factor beta (TGF) could promote hMSCs osteogenic differentiation when the medium contained cAMP-enhancing agents such as IBMX. Identifying key genes, pathways and related miRNAs involved in TGF-induced switch from adipogenic to osteogenic differentiation of hMSCs, is essential to elucidate the molecular mechanisms, find potential drug targets, and identify biomarkers for osteoporosis.

During the past few years, many researchers have turned to using network based approach to get a more comprehensive understanding the process of disease. To better comprehending the relationship between disease-related genes and pathways, this approach can also predict better targets for new agent development. Yang *et al* (4), developed an EgoNet method which can find out the important subnetworks that interrelated to the diseases. This method overcame the problems (the subnetworks are dubious and can not act as a formal topological characteristic) that exist in the methods developed by Dutkowski and Ideker (5) and Zhu *et al* (6). Therefore, we performed EgoNet method (4) to find out the egonetworks associated with osteogenic differentiation of hMSCs.

MicroRNAs (miRNAs) are non-coding single-stranded RNA molecules encoded by endogenous genes of about 20-24 nucleotides in length and involved in post-transcriptional gene expression regulation (7). miRNAs down-regulate the target gene by complementary pairing with the 3'-untranslated regions of these mRNAs. Recently, increasing number of studies are concerning the dysregulation of miRNA in many diseases including osteoporosis (8,9). For example,

microRNA100 (10) and microRNA138 (11) have been proved to function in regulating osteogenic differentiation of hMSCs. Until now, many of the interactions between miRNA and mRNA were found by experimental validation while it is confident but inefficient. However, the systematic computational approach can help us to effective understand the functions of miRNAs by integrating gene expression profiles and miRNA regulatory data (12). Target score algorithm is a probabilistic scoring method which possesses improved prediction accuracy compared to most of the prediction methods (12). Therefore, target score algorithm was performed in this study to identify the miRNAs probably related to the osteogenic differentiation of hMSCs.

In the present study, according to the gene expression profile data and PPI data, we identified the ego-genes by the ego network algorithm. Then we performed KEGG pathway enrichment analysis and predicted the related miRNAs of the ego-genes. The miRNA-mRNA-pathway network was constructed to help comprehensive understanding of the mechanism of hMSCs osteogenic differentiation.

Materials and methods

Data collection. The transcriptomic profiles (GEOD-84500) of hMSCs samples were downloaded from Array Express database (http://www.ebi.ac.uk/arrayexpress/). In the original research, a total of 54 samples were selected for analysis, including 6 control samples and 48 experimental samples. The experimental samples are divided into four groups [BMP2 group (B), BMP2_IBMX group (BI), BMP2_TGFB group (BT), and BMP2_TGFB_IBMX group (BTI)]. The 6 control samples were used as control group for each experimental group. To command the quality of the profile, standard pre-processing was performed (13,14). Furthermore, we converted the expression profile from probe level to gene symbol level, and removed the duplicated symbols. Finally, a total of 20,514 genes were obtained in each group.

All human protein-protein interaction networks (PPIN) were obtained from the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (http://string-db.org/) and used as our global network (include 787,896 interactions). The margin values greater than 0.8 were considered as background PPI. Furthermore, we combined the gene expression data and PPI data by taking the intersections between the two data sets. A total of 48,469 interactions and 7,899 genes were obtained and considered as the background expression network (BEN) applied for the next analysis.

Identification of ego-networks. The Pearson's correlation coefficient (PCC) of the interactions in the above BEN was calculated, and the interactions with absolute PCC >0.8 were chosen to construct the differential expression network (DEN). The four DENs contain 2,361 sides and 1,694 nodes (B), 5,462 sides and 2,382 nodes (BI), 3,520 sides and 1,889 nodes (BT), 4,200 sides and 1,955 nodes (BTI), respectively. We utilized one-side t-test to calculate the p-values of differential gene expression between control and experimental group. Afterwards, the weight value of each interaction was computed using equation 1 where V stands for the set of nodes in the network.

$$W_{ij} = \frac{\sqrt{\log p_i + \log p_j}}{\sqrt{2 * \max l\varepsilon V \left| \log p_l \right|}} \ (if cor(i,j) {\geq} \delta) \ W_{ij} = 0 (if cor(i,j) < \delta)$$

Secondly, all genes in DEN are ranked according to their topological features. The z-score of each gene in the network was calculated using equation 2. Then the z-scores were sorted and the top 5% genes with degree ≥2 were selected as ego genes.

$$g(i) = \sum\nolimits_{j \epsilon Nk(i)} A'_{ijk} g(j)$$

Thirdly, a module search was carried out at the beginning with each ego gene. To ensure accuracy, area under the receiver operating characteristics curve (AUC) was used to evaluate the performance (4). Briefly, for a known ego gene (v), we considered it as module C = {v}. Then the gene u adjacent to v was added to this module and the module C' was obtained. Entropy change was computed as follows: ΔA(C', C) = A (C') - A (C). ΔA (C', C) >0, means the addition of gene u enhanced the classification accuracy of module C. Then, the adjacent genes, which increased ΔA, were added to module C successively until ΔA was no longer increased. Finally, the modules with gene size ≥4 and AUC ≥0.9 were selected as ego-modules.

Pathway enrichment analysis. Kyoto Encyclopedia of Genes and Genomes (KEGG) (15) was performed for functional pathway enrichment of ego-genes. Fisher test was used to identify the enriched pathways. The top 10 pathways with the p-value <0.05 were selected as the differential pathway.

Gene-related microRNA prediction. The microRNAs related to ego-genes were predicted using targetscore algorithm, which was improved based on targetscan method. The TSCS (TargetScan context + score) and PCT (probabilities of conserved targeting) data of the ego-genes were obtained from the TargetScan data website. Based on TSCS and PCT values, VB-GMM algorithm were performed to calculate the target score values of the genes. The target score value was used to screen out the relevant miRNA of ego-genes.

Establishment of miRNA-mRNA-pathway network. The miRNA-mRNA-pathway network was constructed by integrating the mRNA-pathway pairs and miRNA-mRNA pairs we gained in the previous steps. Cytoscape v.3.5.1 software was used for the network construct.

Results

Ego networks. We obtained BEN by taking intersections between the expression profile and PPI data. And the DEN was extracted from the BEN by Pearson correlation coefficient and one-sided t-test. Then the ego-genes were identified from DEN according to the order of their z-scores. The numbers of ego-genes identified in B, BI, BT and BTI groups (ranked in top 5% and degree ≥2) were 84, 119, 94 and 97, respectively. The top 10 ego-genes and their z-scores in each group are shown in Table I. The ego-genes with the highest z-score in B, BI, BT and BTI groups were *DLGAP5*, *DLGAP5*, *NUSAP1* and *NDC80*, respectively.

Table I. Top 10 ego genes of each group extracted from differential expression network.

ID	B		BI		BT		BTI	
	Ego-gene	z-score	Ego-gene	z-score	Ego-gene	z-score	Ego-gene	z-score
1	*DLGAP5*	466.62	*DLGAP5*	1,840.66781	*NUSAP1*	1,684.822609	*NDC80*	1,637.443612
2	*CDCA3*	446.52	*NUSAP1*	1,739.494325	*NCAPG*	1,639.167575	*NUSAP1*	1,527.417318
3	*NUSAP1*	433.23	*TTK*	1,716.473208	*DLGAP5*	1,629.719176	*KIF11*	1,511.460729
4	*TRIP13*	412.35	*NCAPG*	1,658.741346	*CCNB1*	1,617.430721	*DLGAP5*	1,487.441189
5	*AURKA*	409.05	*KIF23*	1,646.915996	*KIF11*	1,596.108219	*TTK*	1,468.464333
6	*KIF23*	369.34	*NDC80*	1,481.402218	*TTK*	1,594.176553	*CCNB1*	1,409.986258
7	*OIP5*	311.49	*KIF11*	1,369.894192	*CDC6*	1,442.414226	*CDC6*	1,350.671948
8	*ANLN*	282.79	*CDCA3*	1,367.5449	*NCAPH*	1,367.570873	*KIF23*	1,236.808446
9	*PBK*	282.79	*OIP5*	1,331.37935	*OIP5*	1,324.601097	*AURKA*	1,207.896451
10	*PRC1*	282.79	*TRIP13*	1,321.589703	*AURKA*	1,317.861436	*NCAPH*	1,195.995795

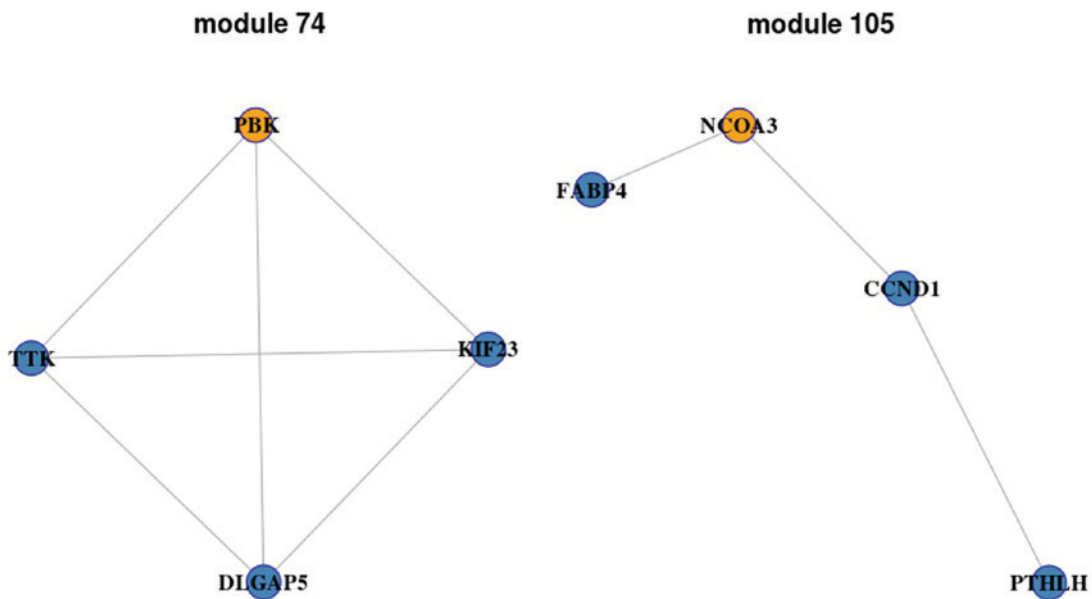B, BMP2 group; BI, BMP2_IBMX group; BT, BMP2_TGFB group; BTI, BMP2_TGFB_IBMX group.

Figure 1. Two significant ego-networks identified by Ego-Net algorithm. Yellow nodes represent ego-genes and blue nodes represent the related genes.

Starting with each ego-gene identified above, we perrformed the module search. The modules with AUC <0.9 were removed and 84, 119, 94 and 97 modules were obtained in B, BI, BT and BTI groups, respectively. Furthermore, the modules with gene size ≥3 were selected as candidate ego-networks. Under this cut-off value, 56 modules in BI group, 1 in BTI group, 2 in B group and none in BT group were obtained. Among these modules, two modules with gene size = 4 were obtained (module 74 and module 105; Fig. 1) in BMP2_IBMX group and none was obtained in the other groups. There are four genes in module 74, including *PBK*, *KIF23*, *TTK* and *DLGAP5*. Module 105 also contains four genes, including *NCOA3*, *FABP4*, *CCND1* and *PTHLH*.

A permutation test was performed on these 59 candidate ego net-works to determine the statistical significance between experimental group and control group. All p-values of these modules were zero. This result indicated that all of them possessed a remarkable difference between experimental group and control group and these modules may be more important than the others in the progression of the switch from adipogenic to osteogenic differentiation of hMSCs.

Enriched pathways of ego-genes. Pathway enrichment analysis was performed based on the KEGG database and Fisher's exact test was used to investigate functional gene sets enriched by ego-genes. Only pathways with adjusted P-value <0.05 were considered to be the differential pathways that the certain ego network enriched. The top 10 enriched pathways of each group are shown in Table II. In BTI group, six ego-genes enriched in Pyrimidine metabolism and six ego-genes enriched in Oocyte meiosis. Four genes were enriched in Fanconi anemia pathway and 3 genes were enriched in Nucleotide excision repair. Six genes were enriched in HTLV-I infection and three genes were enriched in p53 signaling pathway. Two genes were

Table II. The pathways enriched by KEGG analysis of the four groups.

Groups	KEGG Pathway	Adj. p-value	Common gene
B	Proteoglycans in cancer [PATH:hsa05205]	7.28×10^{-6}	CD44, FGF2, FLNC, FN1, HIF1A, ITGAV, ITGB5, PLAUR, PXN
	PI3K-Akt signaling pathway [PATH:hsa04151]	1.47×10^{-5}	CDKN1B, COL1A1, COL1A2, COL2A1, COL4A2, COL5A1, COL6A1, FGF2, FN1, ITGA3, ITGA
	Pathways in cancer [PATH:hsa05200]	5.28×10^{-5}	ADCY4, ADCY7, BIRC5, CDKN1B, COL4A2, FGF2, FN1, HIF1A, ITGA
	Progesterone-mediated oocyte maturation [PATH:hsa04914]	0.000299	ADCY4, ADCY7, BUB1, CCNB1, MAD2L1
	Protein digestion and absorption [PATH:hsa04974]	0.000315	COL1A1, COL1A2, COL4A2, COL5A1, COL6A1
	Small cell lung cancer [PATH:hsa05222]	0.000269	CDKN1B, COL4A2, FN1, ITGA3, ITGAV
	Dilated cardiomyopathy (DCM) [PATH:hsa05414]	0.000332	ADCY4, ADCY7, ITGA3, ITGAV, ITGB5
	Amoebiasis [PATH:hsa05146]	0.000802	COL1A1, COL1A2, COL4A2, COL5A1, FN1
	p53 signaling pathway [PATH:hsa04115]	0.001108	CCNB1, GADD45A, RRM2, SERPINE1
	Platelet activation [PATH:hsa04611]	0.001829	ADCY4, ADCY7, COL1A1, COL1A2, COL5A1
BI	p53 signaling pathway [PATH:hsa04115]	5.64×10^{-6}	CCNB1, CCNB2, CDKN1A, CHEK1, RRM2, SERPINE1
	FoxO signaling pathway [PATH:hsa04068]	2.88×10^{-5}	CCNB1, CCNB2, CDKN1A, CREBBP, PLK4, SMAD3, TGFB1
	Progesterone-mediated oocyte maturation [PATH:hsa04914]	2.52×10^{-5}	ADCY7, BUB1, CCNA2, CCNB1, CCNB2, MAD2L1
	Fanconi anemia pathway [PATH:hsa03460]	0.00043	FANCI, RAD51, RPA3, UBE2T
	Hepatitis B [PATH:hsa05161]	0.00042	BIRC5, CCNA2, CDKN1A, CREBBP, SMAD3, TGFB1
	Pyrimidine metabolism [PATH:hsa00240]	0.00068	PRIM1, RRM1, RRM2, TK1, TYMS
	Pathways in cancer [PATH:hsa05200]	0.00486	ADCY7, BIRC5, CDKN1A, CREBBP, FN1, RAD51, SMAD3, TGFB1
	Colorectal cancer [PATH:hsa05210]	0.008400	BIRC5, SMAD3, TGFB1
	Mismatch repair [PATH:hsa03430]	0.010438	EXO1, RPA3
	5212 Pancreatic cancer [PATH:hsa05212]	0.009972	RAD51, SMAD3, TGFB1
BT	Progesterone-mediated oocyte maturation [PATH:hsa04914]	8.208×10^{-5}	BUB1, CCNA2, CCNB1, CCNB2, MAD2L1
	Mismatch repair [PATH:hsa03430]	0.000213	EXO1, PCNA, RFC4
	HTLV-1 infection [PATH:hsa05166]	0.000320	BUB1B, CCNB2, CDC20, MAD2L1, PCNA, POLE2, PTTG1
	Pyrimidine metabolism [PATH:hsa00240]	0.002032	POLE2, RRM1, RRM2, TYMS
	Nucleotide excision repair [PATH:hsa03420]	0.001793	PCNA, POLE2, RFC4
	p53 signaling pathway [PATH:hsa04115]	0.005144	CCNB1, CCNB2, RRM2
	Base excision repair [PATH:hsa03410]	0.012564	PCNA, POLE2
	Glutathione metabolism [PATH:hsa00480]	0.028618	RRM1, RRM2
	Fanconi anemia pathway [PATH:hsa03460]	0.030731	RAD51, UBE2T
	FoxO signaling pathway [PATH:hsa04068]	0.031864	CCNB1, CCNB2, PLK4
BTI	Pyrimidine metabolism [PATH:hsa00240]	0.000017	POLA1, POLE2, PRIM1, RRM2, TK1, TYMS
	Oocyte meiosis [PATH:hsa04114]	0.000025	AURKA, BUB1, CCNB1, CDC20, FBXO5, MAD2L1
	Fanconi anemia pathway [PATH:hsa03460]	0.000168	BRCA1, FANCI, RAD51, UBE2T
	Nucleotide excision repair [PATH:hsa03420]	0.001942	PCNA, POLE2, RFC4
	HTLV-1 infection [PATH:hsa05166]	0.002318	BUB1B, CCND1, CDC20, MAD2L1, PCNA, POLE2
	p53 signaling pathway [PATH:hsa04115]	0.005559	CCNB1, CCND1, RRM2

Table II. Continued.

Groups	KEGG Pathway	Adj. p-value	Common gene
BTI	Mismatch repair [PATH: hsa03430]	0.006561	*PCNA, RFC4*
	Progesterone-mediated oocyte maturation [PATH: hsa04914]	0.011313	*BUB1, CCNB1, MAD2L1*
	Purine metabolism [PATH: hsa00230]	0.013633	*POLA1, POLE2, PRIM1, RRM2*
	Base excision repair [PATH: hsa03410]	0.013242	*PCNA, POLE2*

B, BMP2 group; BI, BMP2_IBMX group; BT, BMP2_TGFB group; BTI, BMP2_TGFB_IBMX group.

enriched in mismatch repair and three genes were enriched in progesterone-mediated oocyte maturation. Four genes were enriched in Purine metabolism and two genes were enriched in Base excision repair. Most of these ten pathways in BTI group were associated with gene replication and repair which were closely related to cell proliferation and differentiation. The pathways enriched by the ego-genes in the other three group were similar to that in BTI group as described in Table II.

Prediction of miRNAs related to ego-genes. The miRNAs related to ego-genes were predicted using Target Score algorithm. The relevant miRNAs with a target score value >0.4 were selected. There were 423 pairs in B group, 570 pairs in BI group, 433 pairs in BT group and 456 miRNA-mRNA pairs in BTI group. Relationships between ego-genes of BI group and the predicted microRNAs are shown in Fig. 2. Eleven mRNAs in BI group (*NCOA3*, *ANLN*, *CHEK1*, *KIF23*, *RACGAP1*, *TRIP13*, *CDKN1A*, *CDC6*, *PBK*, *NCAPG* and *FN1*) were known to be targets of 1-5 miRNAs. The relationships between ego-genes of other groups and the predicted microRNAs were also constructed (data not shown). There are 9 mRNAs in B group (*CDKN1B*, *ANLN*, *WEE1*, *KIF23*, *TRIP13*, *HIF1A*, *PBK*, *FN1* and *FGF2*) known to be targets of 1-3 miRNAs. Seven mRNAs (*ANLN*, *KIF23*, *RACGAP1*, *TRIP13*, *CDC6*, *PBK* and *NCAPG*) in group BT have been validated be the targets of 1 miRNA. Seven mRNAs (*ANLN*, *KIF23*, *TRIP13*, *CCND1*, *CDC6*, *PBK* and *NCAPG*) of these pairs in BTI group were known to be targets of 1-3 miRNAs.

Construction and analysis of miRNA-mRNA-pathway network. We obtained 267 mRNA-pathway pairs and 423 miRNA-mRNA pairs of the ego-genes in B group; 230 mRNA-pathway pairs and 570 miRNA-mRNA pairs of the ego-genes in BI group; 94 mRNA-pathway pairs and 433 miRNA-mRNA pairs of the ego-genes in BT group; 117 mRNA-pathway pairs and 456 miRNA-mRNA pairs of the ego-genes in BTI group. The miRNA-mRNA-pathway networks were constructed by merging these pairs. The miRNA-mRNA-pathway network we constructed for BI group is shown in Fig. 3. This network not only reflects the regulatory association between miRNA and mRNA, but also suggests the biological function of miRNA regulation. The networks of other 3 groups were also constructed (data not shown).

Discussion

Over the past 10 years, considerable work has been done on predicting gene markers from expression profiles, which can classify and give prognosis of the disease, but these predictions often lack reliability and accuracy (16). Therefore, researchers need to turn to gain a more comprehensive understanding of the disease by acquiring subnetworks. In the present study, we used EgoNet algorithm proposed by Yang *et al* (4), to identify the ego-networks associated with the switch from adipogenic to osteogenic differentiation of hMSCs. The pathway enrichment analysis and related miRNAs prediction were performed by KEGG and targetscore, respectively. Finally the miRNA-mRNA-pathway networks were constructed.

Through EgoNet algorithm method, we found that Nucleolar and spindle associated protein 1 (NUSAP1) and Discs, large
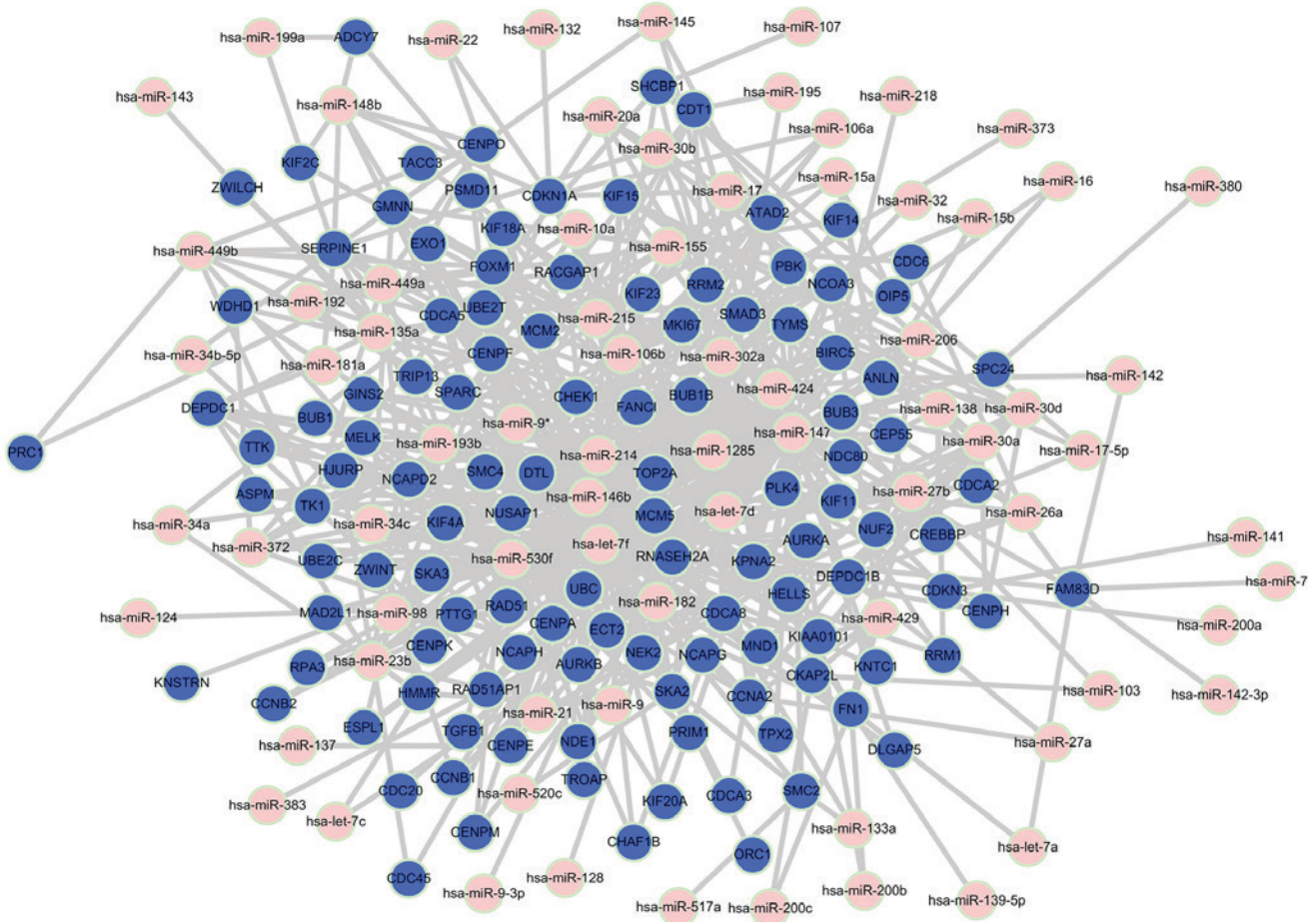
Figure 2. Relationships between ego-genes of BI group and the predicted microRNAs. The pink and blue nodes in the network represented miRNA and mRNA, respectively. The edges represent the relationships between miRNA, and mRNA.

(*Drosophila*) homolog-associated protein 5 (DLGAP5) were two genes that with higher z-scores and ranked in front in all the four groups. NUSAP1 is a regulatory cyclin, which take a vital part in the spindle assembly and cell proliferation. NUSAP1 has been shown to be overexpressed in many tumors, so its functional disorders may be closely related to the development of tumors (17). So far, there is no report on the relationship between NUSAP1 and osteoporosis. In the present study, the abnormality of NUSAP1 compared to the control group may be related to the differentiation of hMSCs. DLGAP5 is a protein involved in the formation of centromere in cell division and played an important role in the movement of spindle (18). DLGAP5 has also been shown to be involved in many cancer types, including breast cancer, prostate cancer and liver cancer (18-20). Nuclear division cycle 80 (NDC80) gene gained the highest z-score in BTI group (Table I). It is a nuclear protein which is closely related to the stability of chromosomes and takes important part in mitosis regulation. It has been known that the irregular expression of NDC80 in cells often accompanied with spindle checkpoint abnormal, chromosome separation abnormal and cell cycle disorders, which may eventually lead to the occurrence of cancer (21). Further study is needed to verify whether NUSAP1, DLGAP5 and NDC80 are indeed associated with the differentiation of hMSCs.

Through module search, we found two modules with gene size = 4 in BI group. The corresponding ego-genes were PDZ binding kinase (PBK) and Nuclear receptor coactivator 3 (NCOA3). PBK is a serine/threonine protein kinase which is associated with the mitogen-activated protein kinase (MAPK) kinase (MAPKK) family (22). It has been reported by many studies that overexpression of PBK may be involved in tumorigenesis (23,24). NCOA3 is a number of the steroid receptor coactivator p160 family, which regulates the transcription of nuclear receptors activation through the combination of nuclear receptors. It was found that the mouse model of NCOA3 knockout showed a decrease in lipid fat deposition and an increase in energy consumption (25-27). Hartig *et al* (28), reported that NCOA3 can combine with peroxisome proliferator activated receptor gamma (PPARγ) and lead to a decrease in PPARγ-s114 phosphorylation. Furthermore, it promoted the nuclear receptor PPARγ transcription activity and thus promoted the deposition of fat. It can be seen that NCOA3 is closely related to fat deposition. In the present study, the different expression of NCOA3 in hMSCs may be related to switch from adipogenic to osteogenic differentiation of hMSCs.

A miRNA-mRNA-pathway network was constructed in each group by combining the miRNA-mRNA pairs predicted by targetscore algorithm and the mRNA-pathway pairs together.
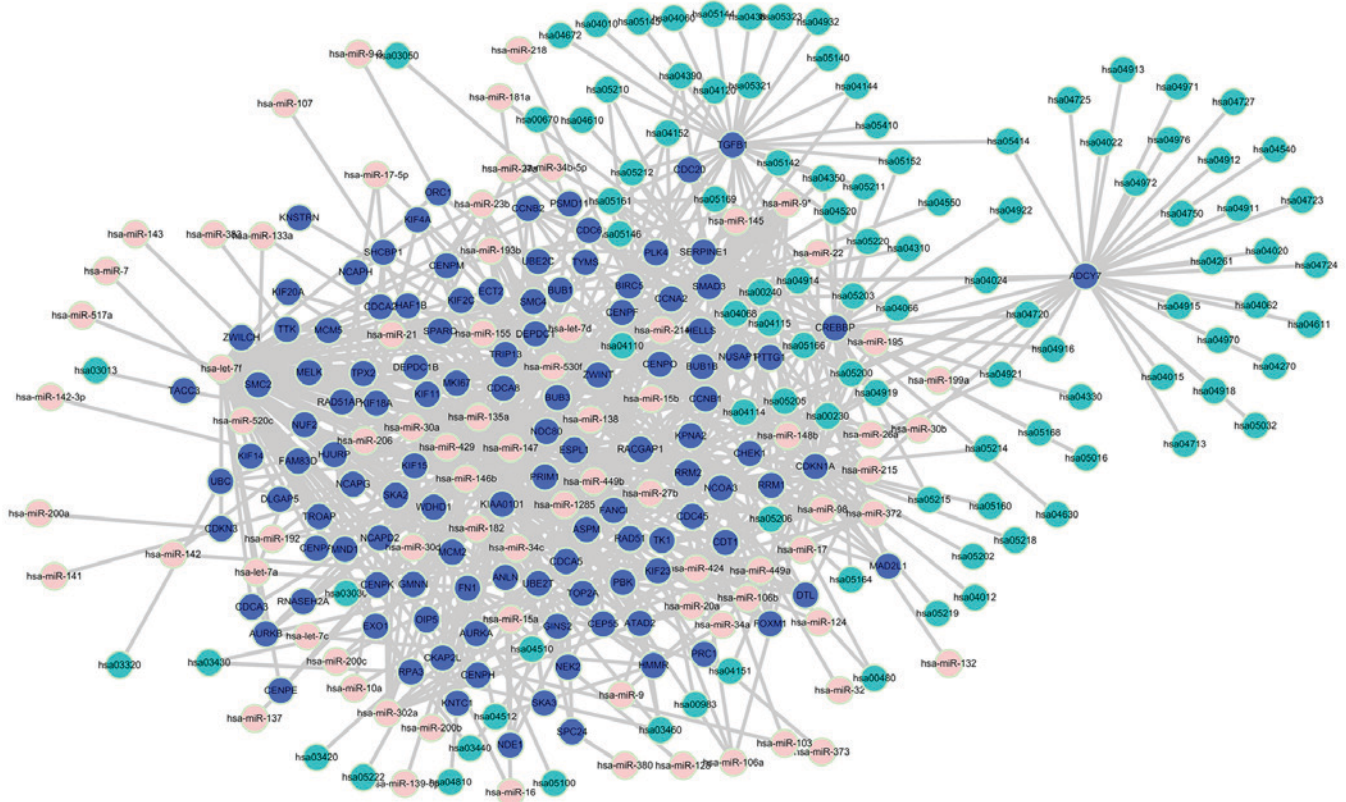
Figure 3. miRNA-mRNA-pathway network of group BI. The pink, blue and green nodes in the network represent miRNA, mRNA, pathway, respectively. The edges represent the relationships between miRNA, mRNA and pathways. miRNA, microRNA; BI, BMP2_IBMX.

Among the network constructed in BI group, CDKN1A is involved in four of the top 10 enriched pathways including p53 signaling pathway, FoxO signaling pathway, Hepatitis B, and pathways in cancer. CDKN1A was predicted to be the target of 10 miRNAs (hsa-miR-98, *hsa-miR-20a*, *hsa-miR-106b*, *hsa-miR-106a*, hsa-miR-22, *hsa-miR-17*, *hsa-miR-132*, hsa-miR-148b, hsa-miR-1285, hsa-let-7d) and five pairs of them have been validated (italicized miRNAs) (29-31). Cell cycle inhibitor p21 CDKN1A is widely known as a cell cycle inhibitor. Besides it is also involved in cell motility, cell death program, DNA replication and repair, and transcription (32). Therefore, CDKN1A may be one of the key genes involved in hMSCs differentiation.

Above all, we identified the differential gene expression networks between the control and experimental groups, and identified the ego-genes by ego-net analysis. Subsequently, the key pathways and related miRNA were identified through the KEGG pathway enrichment analysis and TargetScore algorithm. Finally, the mRNA-miRNA-pathway network was constructed. These identified key genes and pathways gave a deeper understanding of the molecular mechanisms underlying the transformation of TGFβ-induced adipocytes into osteoblasts. Furthermore, the constructed network gave us a more comprehensive understanding of the mechanism of osteogenic differentiation of hMSCs. However, the relationships between the identified miRNA, mRNA and hMSC differentiation still need to be validated by biological experiments.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

WXP conceived the study and wrote the manuscript. WXP and CHG acquired and analyzed the data. GBH assisted in

the explanation of the results and revised the manuscript. All authors read and approved the final manuscript.

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. Cosman F, de Beur SJ, LeBoff MS, Lewiecki EM, Tanner B, Randall S and Lindsay R: Clinician's guide to prevention and treatment of osteoporosis. Osteoporos Int 26: 2045-23047, 2015.
2. Rosen CJ, Ackert-Bicknell C, Rodriguez JP and Pino AM: Marrow fat and the bone microenvironment: Developmental, functional, and pathological implications. Crit Rev Eukaryot Gene Expr 19: 109-124, 2009.
3. Justesen J, Stenderup K, Ebbesen EN, Mosekilde L, Steiniche T and Kassem M: Adipocyte tissue volume in bone marrow is increased with aging and in patients with osteoporosis. Biogerontology 2: 165-171, 2001.
4. Yang R, Bai Y, Qin Z and Yu T: EgoNet: Identification of human disease ego-network modules. BMC Genomics 15: 314, 2014.
5. Dutkowski J and Ideker T: Protein networks as logic functions in development and cancer. PLOS Comput Biol 7: e1002180, 2011.
6. Zhu Y, Shen X and Pan W: Network-based support vector machine for classification of microarray samples. BMC Bioinformatics 10 (Suppl 1): S21, 2009.
7. Khan AA, Betel D, Miller ML, Sander C, Leslie CS and Marks DS: Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs. Nat Biotechnol 27: 549-555, 2009.
8. Lerebours F, Cizeron-Clairac G, Susini A, Vacher S, Mouret-Fourme E, Belichard C, Brain E, Alberini JL, Spyratos F, Lidereau R, *et al*: miRNA expression profiling of inflammatory breast cancer identifies a 5-miRNA signature predictive of breast tumor aggressiveness. Int J Cancer 133: 1614-1623, 2013.
9. Pignot G, Cizeron-Clairac G, Vacher S, Susini A, Tozlu S, Vieillefond A, Zerbib M, Lidereau R, Debre B, Amsellem-Ouazana D, *et al*: microRNA expression profile in a large series of bladder tumors: Identification of a 3-miRNA signature associated with aggressiveness of muscle-invasive bladder cancer. Int J Cancer 132: 2479-2491, 2013.
10. Zeng Y, Qu X, Li H, Huang S, Wang S, Xu Q, Lin R, Han Q, Li J and Zhao RC: MicroRNA-100 regulates osteogenic differentiation of human adipose-derived mesenchymal stem cells by targeting BMPR2. FEBS Lett 586: 2375-2381, 2012.
11. Eskildsen T, Taipaleenmäki H, Stenvang J, Abdallah BM, Ditzel N, Nossent AY, Bak M, Kauppinen S and Kassem M: MicroRNA-138 regulates osteogenic differentiation of human stromal (mesenchymal) stem cells in vivo. Proc Natl Acad Sci USA 108: 6139-6144, 2011.
12. Li Y, Goldenberg A, Wong KC and Zhang Z: A probabilistic approach to explore human miRNA targetome by integrating miRNA-overexpression data and sequence information. Bioinformatics 30: 621-628, 2014.
13. Bolstad BM, Irizarry RA, Astrand M and Speed TP: A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19: 185-193, 2003.
14. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B and Speed TP: Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 31: e15, 2003.
15. Aoki-Kinoshita KF and Kanehisa M: Gene annotation and pathway mapping in KEGG. Methods Mol Biol 396: 71-91, 2007.
16. Venet D, Dumont JE and Detours V: Most random gene expression signatures are significantly associated with breast cancer outcome. PLOS Comput Biol 7: e1002240, 2011.
17. Gordon CA, Gong X, Ganesh D and Brooks JD: NUSAP1 promotes invasion and metastasis of prostate cancer. Oncotarget 8: 29935-29950, 2017.
18. Liao W, Liu W, Yuan Q, Liu X, Ou Y, He S, Yuan S, Qin L, Chen Q, Nong K, *et al*: Silencing of DLGAP5 by siRNA significantly inhibits the proliferation and invasion of hepatocellular carcinoma cells. PLoS One 8: e80789, 2013.
19. Liu R, Guo CX and Zhou HH: Network-based approach to identify prognostic biomarkers for estrogen receptor-positive breast cancer treatment with tamoxifen. Cancer Biol Ther 16: 317-324, 2015.
20. Gomez CR, Kosari F, Munz JM, Schreiber CA, Knutson GJ, Ida CM, El Khattouti A, Karnes RJ, Cheville JC, Vasmatzis G, *et al*: Prognostic value of discs large homolog 7 transcript levels in prostate cancer. PLoS One 8: e82833, 2013.
21. Xing XK, Wu HY, Chen HL and Feng HG: NDC80 promotes proliferation and metastasis of colon cancer cells. Genet Mol Res 15: gmr.15028312, 2016.
22. Dougherty JD, Garcia AD, Nakano I, Livingstone M, Norris B, Polakiewicz R, Wexler EM, Sofroniew MV, Kornblum HI and Geschwind DH: PBK/TOPK, a proliferating neural progenitor-specific mitogen-activated protein kinase kinase. J Neurosci 25: 10773-10785, 2005.
23. Yang YF, Pan YH, Cao Y, Fu J, Yang X, Zhang MF and Tian QH: PDZ binding kinase, regulated by FoxM1, enhances malignant phenotype via activation of β-Catenin signaling in hepatocellular carcinoma. Oncotarget 8: 47195-47205, 2017.
24. Shih MC, Chen JY, Wu YC, Jan YH, Yang BM, Lu PJ, Cheng HC, Huang MS, Yang CJ, Hsiao M, *et al*: TOPK/PBK promotes cell migration via modulation of the PI3K/PTEN/AKT pathway and is associated with poor prognosis in lung cancer. Oncogene 31: 2389-2400, 2012.
25. Louet JF, Coste A, Amazit L, Tannour-Louet M, Wu RC, Tsai SY, Tsai MJ, Auwerx J and O'Malley BW: Oncogenic steroid receptor coactivator-3 is a key regulator of the white adipogenic program. Proc Natl Acad Sci USA 103: 17868-17873, 2006.
26. Ma X, Xu L, Wang S, Cui B, Li X, Xu J and Ning G: Deletion of steroid receptor coactivator-3 gene ameliorates hepatic steatosis. J Hepatol 55: 445-452, 2011.
27. Coste A, Louet JF, Lagouge M, Lerin C, Antal MC, Meziane H, Schoonjans K, Puigserver P, O'Malley BW and Auwerx J: The genetic ablation of SRC-3 protects against obesity and improves insulin sensitivity by reducing the acetylation of PGC-1{alpha}. Proc Natl Acad Sci USA 105: 17187-17192, 2008.
28. Hartig SM, He B, Long W, Buehrer BM and Mancini MA: Homeostatic levels of SRC-2 and SRC-3 promote early human adipogenesis. J Cell Biol 192: 55-67, 2011.
29. Ivanovska I, Ball AS, Diaz RL, Magnus JF, Kibukawa M, Schelter JM, Kobayashi SV, Lim L, Burchard J, Jackson AL, *et al*: MicroRNAs in the miR-106b family regulate p21/CDKN1A and promote cell cycle progression. Mol Cell Biol 28: 2167-2174, 2008.
30. Ambs S, Prueitt RL, Yi M, Hudson RS, Howe TM, Petrocca F, Wallace TA, Liu CG, Volinia S, Calin GA, *et al*: Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer. Cancer Res 68: 6162-6170, 2008.
31. Wu S, Huang S, Ding J, Zhao Y, Liang L, Liu T, Zhan R and He X: Multiple microRNAs modulate p21Cip1/Waf1 expression by directly targeting its 3' untranslated region. Oncogene 29: 2302-2308, 2010.
32. Dutto I, Tillhon M, Cazzalini O, Stivala LA and Prosperi E: Biology of the cell cycle inhibitor p21(CDKN1A): Molecular mechanisms and relevance in chemical toxicology. Arch Toxicol 89: 155-178, 2015.