Cell and Tissue Research (2019) 378:513–529
https://doi.org/10.1007/s00441-019-03071-1

REGULAR ARTICLE

Co-expression network analysis identified key genes in association with mesenchymal stem cell osteogenic differentiation

Wang Yang[1,2] · Yuhan Xia[1] · Xiaoli Qian[1] · Meijing Wang[1] · Xiaoling Zhang[3,4] · Yulin Li[1] · Lisha Li[1]

Received: 22 September 2018 / Accepted: 3 July 2019 / Published online: 15 August 2019

Abstract

Although several studies have shown that osteogenic differentiation of different mesenchymal stem cell (MSC) lines can be guided by the 3D scaffold with growth factors or biochemical agent, the key mechanism regulating osteogenic differentiation is not known yet. Here, this study was designed to investigate key genes that regulate the induction of osteogenesis by different MSC lines in different ways. Expression profiling by array (GSE58919 and GSE18043) was downloaded and analyzed using weighted gene co-expression network analysis (WGCNA) to narrow genes associated with osteogenic differentiation. A protein-protein interactive (PPI) network was built to find the key genes and the role of these key genes was confirmed by statistical analysis. To understand the function of genes associated with osteogenesis, gene ontology (GO) and the Kyoto encyclopedia of genes and genomes (KEGG) were analyzed, which showed that key genes in MSC osteogenic differentiation induced by a biochemical agent involve regulation of cell apoptosis and proliferation while key genes in MSC osteogenic differentiation induced by the 3D scaffold with growth factors involve regulation of cajal body and centromeres. Furthermore, 58 key genes are involved in Wnt signaling pathway, ion response and focal adhesion. Proteasome also played a key role in osteogenic differentiation. Seven potential key genes were found essential in the osteogenic differentiation of MSCs in the PPI network, especially the five key genes, CCT2, NOP58, FBL, EXOSC8 and SNRPD1. This study will provide important targets of MSC osteogenic differentiation that will help us understand the mechanism of osteogenic differentiation in MSCs.

Keywords Mesenchymal stem cells · WGCNA · Osteogenic differentiation · Different cultural method · GEO

Introduction

In recent years, the mesenchymal stem cell (MSC) has become a key cell source in regenerative medicine and other fields. This is due to the wide range of MSC sources, including fat, bone marrow and umbilical cord. In addition, MSC, as a pluripotent stem cell, can be directionally induced into the osteoblast, chondrocyte (Chen et al. 2018; Shuai et al. 2018), myocyte (Sassoli et al. 2018), neuroblast (Yan et al. 2018), and so on. The development of these basic studies has made it possible for stem cell therapy to progress to clinical practice in the future (Li et al. 2018; Yan et al. 2018). Due to the broad prospect of stem cell therapy (Golpanian et al. 2016; Huebsch et al. 2015; Mathiasen et al. 2015), the mechanism of induced osteogenic differentiation of MSCs has attracted much attention.

Based on the in-depth study of MSC differentiation, varieties of methods were used to induce MSC differentiation. Three-dimensional (3D) scaffold with growth factors (McMillan et al. 2018) or a biochemical agent is a common

Wang Yang and Yuhan Xia contributed equally to this study.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00441-019-03071-1) contains supplementary material, which is available to authorized users.

✉ Xiaoling Zhang
zhangxl714@sina.com

✉ Yulin Li
ylli@jlu.edu.cn

✉ Lisha Li
43973966@qq.com

[1] The Key Laboratory of Pathobiology, Ministry of Education, Norman Bethune College of Medicine, Jilin University, Changchun 130021, China

[2] College of Clinical Medicine, Jilin University, Changchun, China

[3] The First Hospital, Jilin University, Changchun 130061, China

[4] Institute of Immunology, Jilin University, Changchun 130061, China

method used to induce MSC directional differentiation. Multiple methods can induce the MSCs to differentiate into the same cell line, which means that there is a common path in these different ways to regulate MSC directional differentiation.

Gene expression omnibus (GEO), a National Center for Biotechnology Information database for gene expression, can provide a large number of high-throughput gene expression data hybridization arrays, chips and microarrays for our analysis. Data gained from the GEO database to analyze can provide some forward-looking ideas. As the public database continues to grow, the corresponding analysis methods are also changing. Weighted gene co-expression network analysis (WGCNA), as an excellent analytical method, which can obtain the correlation between the traits and the co-expression gene modules by constructing scale-free networks, has been widely used to find hub genes related to traits.

In this study, we used WGCNA to analyze RNA expression data that we obtained, which was produced by the osteogenic differentiation of MSC in different common ways, so as to explore the genes that play a key role in the process of MSC osteogenic differentiation induced by different treatments.

Materials and methods

RNA expression data processing

The RNA expression data and trait data of MSC-induced differentiation were downloaded from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/). Database GSE58919 was performed on [HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version] and database GSE18043 was performed on [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array, which are used to construct co-expression networks by WGCNA and identify the key genes in this research. In GSE58919, MSCs were induced fibrillization and osteogenic differentiation by 3D scaffold with specific growth factors, while MSCs were induced osteogenic differentiation by dexamethasone in GSE18043 (Hamidouche et al. 2009).

Co-expression network construction

The matrix data downloaded from the GEO database were treated with standardized processing. Before the construction of the co-expression network, we filtered the data. Median absolute deviation (MAD) is a robust measure of sample diversity in a univariate dataset. Following the requirements, we removed three expression data forms of GSE58919, namely GSM1422303, GSM1422304 and GSM1422305. GSE58919 comes from the induction of 3D scaffold with different growth factors in MSCs, while GSE18043 is derived from induction of osteo-inductive fluid in MSCs. Using MAD calculations, we selected the top 5000 genes from the database of GSE58919 with 18,992 genes and the database of GSE18043 with 20,848 genes. Then, we used the WGCNA package (Langfelder and Horvath 2008) in R to construct scale-free co-expression network for the top 5000 genes accordingly.

In the first step, we computed the person's coefficient between any two genes. To determine whether two genes share similar patterns of expression, the weighted value of correlation coefficient in WGCNA analysis makes the connections among genes in a co-expression network that obeys scale-free networks, which is more biologically significant.

In the second step, a hierarchical clustering tree was constructed from the correlation coefficient between genes. Different branches of the clustering tree represent different gene modules and different colors represent different modules. Based on the weighted correlation coefficient of a gene, the gene was classified according to the expression pattern and the pattern similarity gene was classified as a module.

In this study, the soft threshold in GSE68919 is 16 while the soft threshold in GSE18043 is 14. We selected a minimum size (gene module) of 100 and two modules were merged with more than 75% similarity. It is important to note that the gray module refers to a collection of genes that cannot be aggregated into other modules.

Screening key modules

Through the association analysis of each module and the sample feature vector, modules that are significantly related to MSC differentiation can be selected and hub genes affecting MSC differentiation are selected from the selected modules. We calculated the gene significance (GS) value of the gene in the module by reckoning the correlation between the module eigenvector gene (ME) of the network module and the sample feature vector. Then, the correlation of GS and module membership (MM) was used to test whether the size of MM was related to the differentiation of MSC.

Venn diagram of key genes

Through the WGCNA analysis, we obtained the relationship diagram of the trait and the module. The blue module in database GSE18043 and the brown and blue modules in database GSE58919 are involved in maintaining the differentiation of MSC. Then, we drew a Venn diagram to sort out the parts that are common to both databases GSE18043 and GSE58919. Venny v2.1.0 (http://bioinfogp.cnb.csic.es/tools/venny/) was used to draw a Venn diagram.

Table 1 The merged colors and number of genes from the obtained modules in GSE58919 and GSE18043

Merged colors	Number of genes
GSE58919	
Blue	595
Brown	559
Green	3087
Green yellow	192
Gray	184
Yellow	383
GSE18043	
Black	300
Blue	563
Brown	2065
Cyan	160
Green	395
Gray	63
Magenta	901
Midnight blue	158
Salmon	185
Tan	210

GO analysis and KEGG pathway analysis

Through the Venn diagram, we obtained the genes that are common to the modules and the genes that are specific to each module. We performed gene ontology (GO) analysis and Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis of GSE58919 blue, GSE58919 brown, GSE18043 blue-specific, GSE18043 blue–GSE58919 brown and GSE18043 blue–GSE58919 blue genes. It is worth mentioning that the reason why we combined GSE18043 blue–GSE58919 brown and GSE18043 blue–GSE58919 blue intersectional genes was based on the WGCNA analysis. There is a positive correlation between brown and blue modules in GSE58919, indicating that there is not much difference between the two blocks and the combination is beneficial to subsequent analysis.

Through Enrichr (http://amp.pharm.mssm.edu/Enrichr/) (Chen et al. 2013; Kuleshov et al. 2016), we obtained the GO term and KEGG pathway results of the above-mentioned grouping genes. Both the GO biological process and KEGG pathway determined the significance that was based on the combined score. The top 10 terms in the GO analysis and pathways in the KEGG pathway were drawn by a histogram and Cytoscape v3.6.0 (Killcoyne et al. 2009; Shannon et al. 2003).

PPI network

We also used STRING database v10 (Szklarczyk et al. 2015) to construct a protein–protein interactive (PPI) network between GSE18043 blue–GSE58919 brown and GSE18043 blue–GSE58919 blue combined genes. We showed the PPI network through Cytoscape v3.6.0. The importance of each node is measured by the degree of each node in the PPI network constructed.

Heatmap of key gene

We used R to draw the heatmap of 58 key genes selected in both GSE58919 and GSE18043.

Statistics of key gene selected in PPI network

The key nodes picked out from the PPI network were displayed using a boxplot to find a significantly different expression of the data of GSE58919.

Gene set enrichment analysis of key gene

To verify the biological function of the key gene, we used their expression levels to classify the samples into high and low expression of the key gene group. The KEGG pathway was then enriched using the Broad Institute's gene set enrichment analysis (GSEA) software (Mootha et al. 2003; Subramanian et al. 2005). The mRNA data of the top 5000 MAD in GSE58919 were used for GSEA. The parameters of the software were default. The normalized enrichment score (NES) value and the p value of the enriched pathways were used to confirm the accuracy of the results. $|NES| > 1$ and p value < 0.05 were used to filter the pathways. The pathways with $|NES| > 1$, p value < 0.05 and FDR < 0.25 were considered as meaningful. Then, we used "ggplot2" and "ggrepel" packages to show the results in the R platform.

Results

RNA expression data acquisition and data filtration

After obtaining RNA expression data and quality assessment, the expression matrices were obtained from 21 samples in GSE58919, in which 3 expression data were removed, namely GSM1422303, GSM1422304 and GSM1422305. Another expression matrix was obtained from 12 samples in GSE18043. The top 5000 genes of each dataset under the threshold of MAD were selected for subsequent WGCNA analysis.

Co-expression network construction and key module screening

Eighteen samples in GSE58919 and 12 samples in GSE18043 with trait data were obtained in co-expression analysis (Fig.

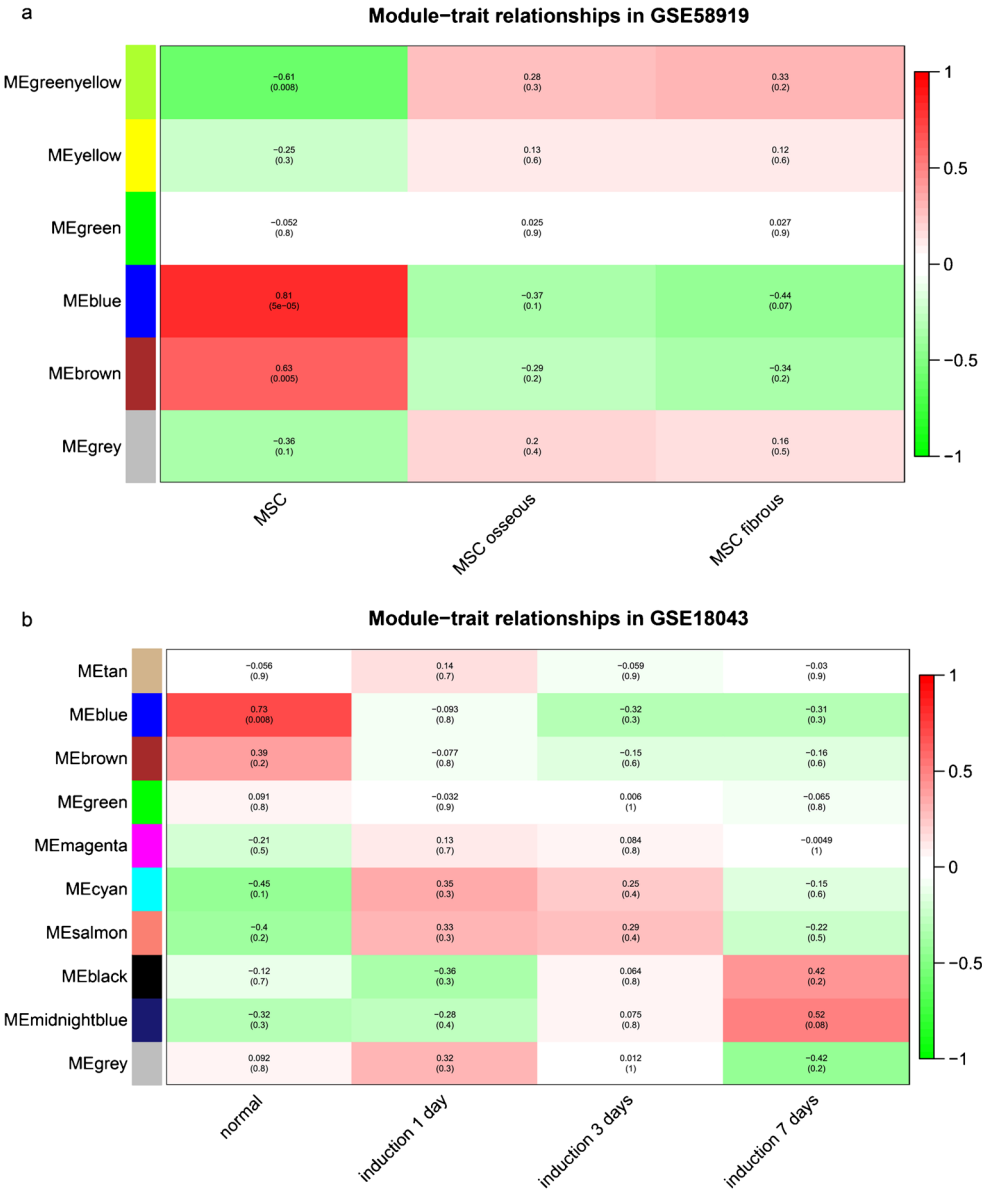


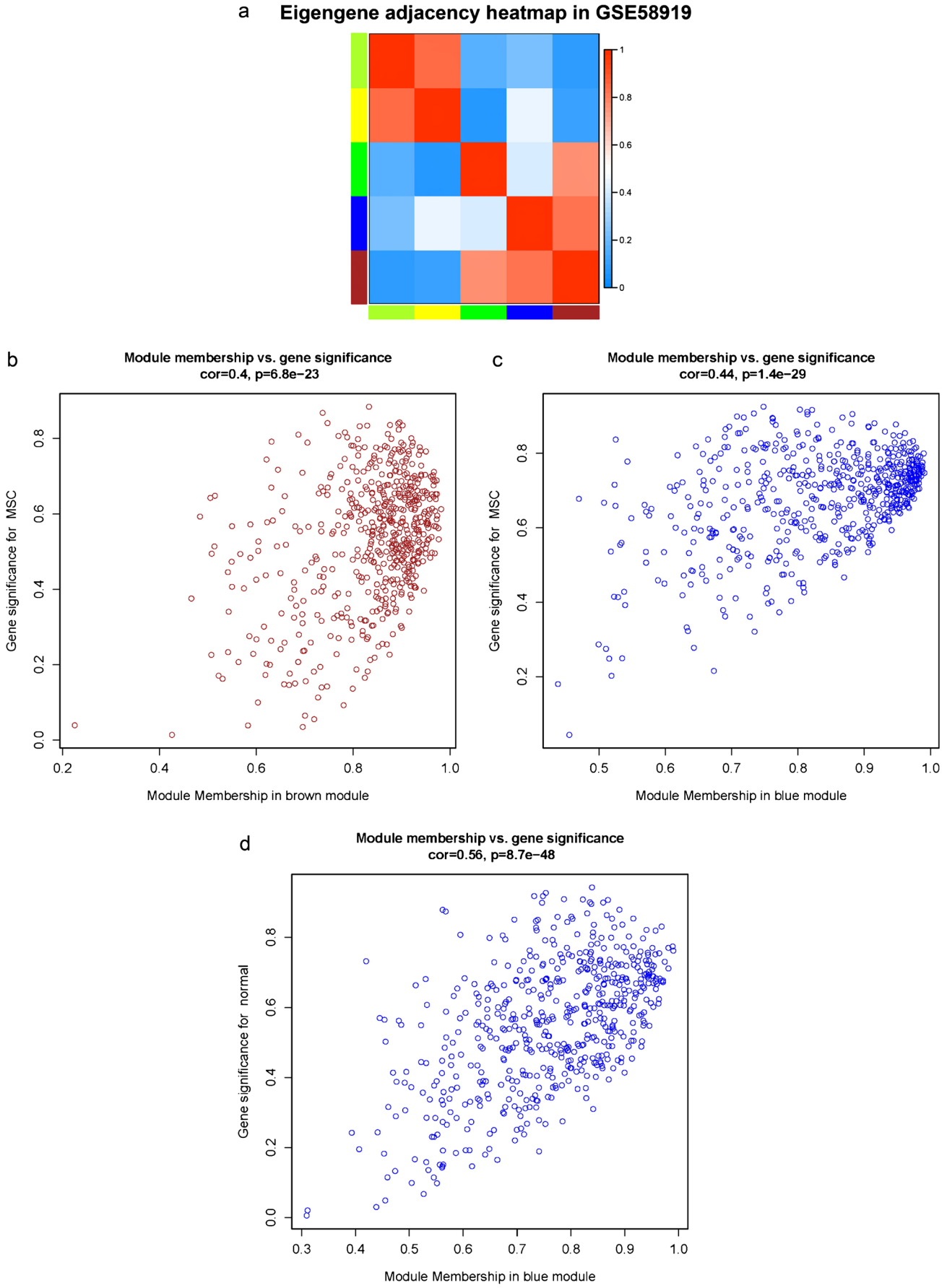
Fig. 1 (**a**) Module–trait relationship of GSE58919. (**b**) Module–trait relationship of GSE18043. The first row of each cube represents the correlation coefficient between the module and the trait and the second row represents the significance of the correlation coefficient. The color of the cube is measured by correlation coefficient; red represents positive correlation and blue represents negative correlation

S1a,b). In GSE58919, we selected WGCNA best soft-thresholding $\beta = 16$ to ensure a scale-free network while $\beta = 14$ was selectedin GSE18043 (Fig. S2). In GSE18043, we acquired 10 modules, while 6 modules were obtained in GSE58919 (Table 1, Fig. S3). In the module–trait relationship obtained, we found modules related to the osteogenic differentiation of MSC (blue and brown modules in GSE58919 and blue module in GSE18043), which were explored in both datasets (Fig. 1).

Identification of key genes for MSC osteogenic differentiation in key modules

After obtaining the key modules, brown and blue modules existed in GSE58919. Through the eigengene adjacency heatmap, there was positive correlation between brown and blue modules in GSE58919, suggesting these two modules show two similar processes and are both involved in the osteogenic differentiation of MSC. We obtained module membership (MM) vs gene significance (GS) of GSE58919 brown, GSE58919 blue and GSE18043 blue modules (Fig. 2). Then, we used the Venn diagram to obtain coincident genes in the key modules selected in the two datasets, which were regarded as the key genes to regulate osteogenic differentiation of MSCs, including 33 genes in both GSE18043 blue and GSE58919 brown modules and 25 genes in both GSE18043 blue and GSE58919 blue modules. A total of 58 key genes were acquired by overlapping key modules

Fig. 2 Eigengene adjacency heatmap and module membership vs. gene significance. (**a**) Eigengene adjacency heatmap in GSE58919. (**b**–**d**) Module membership vs. gene significance in GSE58919 brown, GSE58919 blue and GSE18043 blue modules, respectively. Brown module and blue module are correlated in GSE58919 ▶


a
Eigengene adjacency heatmap in GSE58919
1
0.8
0.6
0.4
0.2
0
b
Module membership vs. gene significance
cor=0.4, p=6.8e−23
Gene significance for MSC
Module Membership in brown module
c
Module membership vs. gene significance
cor=0.44, p=1.4e−29
Gene significance for MSC
Module Membership in blue module
d
Module membership vs. gene significance
cor=0.56, p=8.7e−48
Gene significance for normal
Module Membership in blue module

(Fig. 3a,b). Through the heatmap, we found that in 3D culture with growth factors, two cell lines (MSC and iPS-MSC) had the same trend of osteogenic differentiation (Fig. 4a). These 58 genes also showed similar trends in the cells differentiated by osteoinductive fluid alone (Fig. 4b).

GO analysis and KEGG pathway analysis

By analyzing the GO and KEGG pathway enrichments of the three modules of GSE58919 blue, GSE58919 brown and GSE18043 blue and their key genes, as a result, only top the 10 GO terms or KEGG pathways were displayed. Pathway enrichments in GSE58919 were used to display the main pathway induced by 3D scaffold with growth factors. In the enrichment of the GSE58919 brown module, most of the processes were focused on the regulation of the cajal body and the degradation process with the involvement of ubiquitination while the most significant KEGG pathway was involved in proteasome. In the enrichment of the GSE58919 blue module, the top 10 GO terms were referred to as the functional

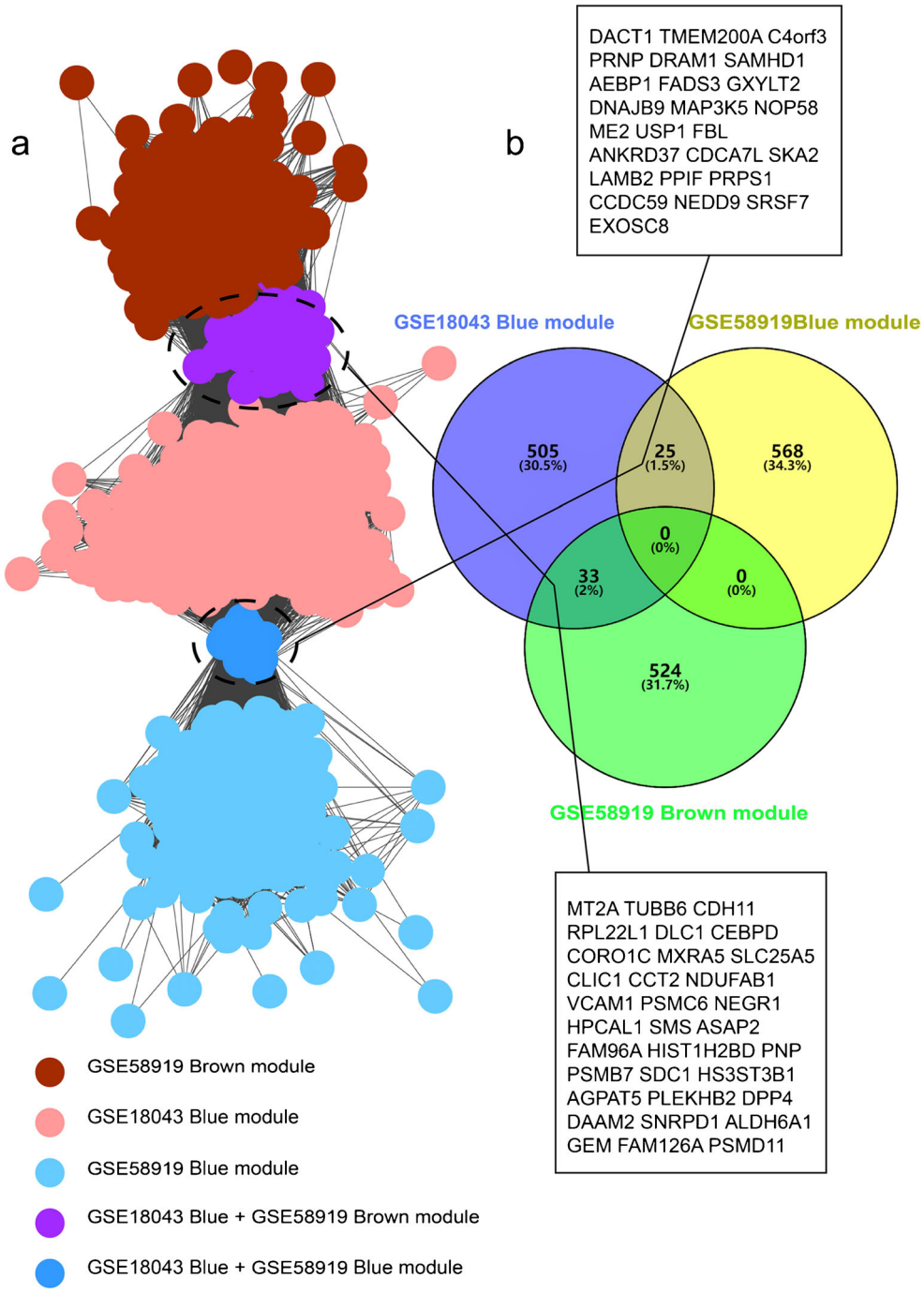


Fig. 3 Key module gene network and Venn diagram for seeking key genes. (**a**) Distribution of module-specific or non-module-specific genes among GSE58919 blue, GSE58919 brown and GSE18043 blue modules. (**b**) Venn diagram of different modules and key genes selected

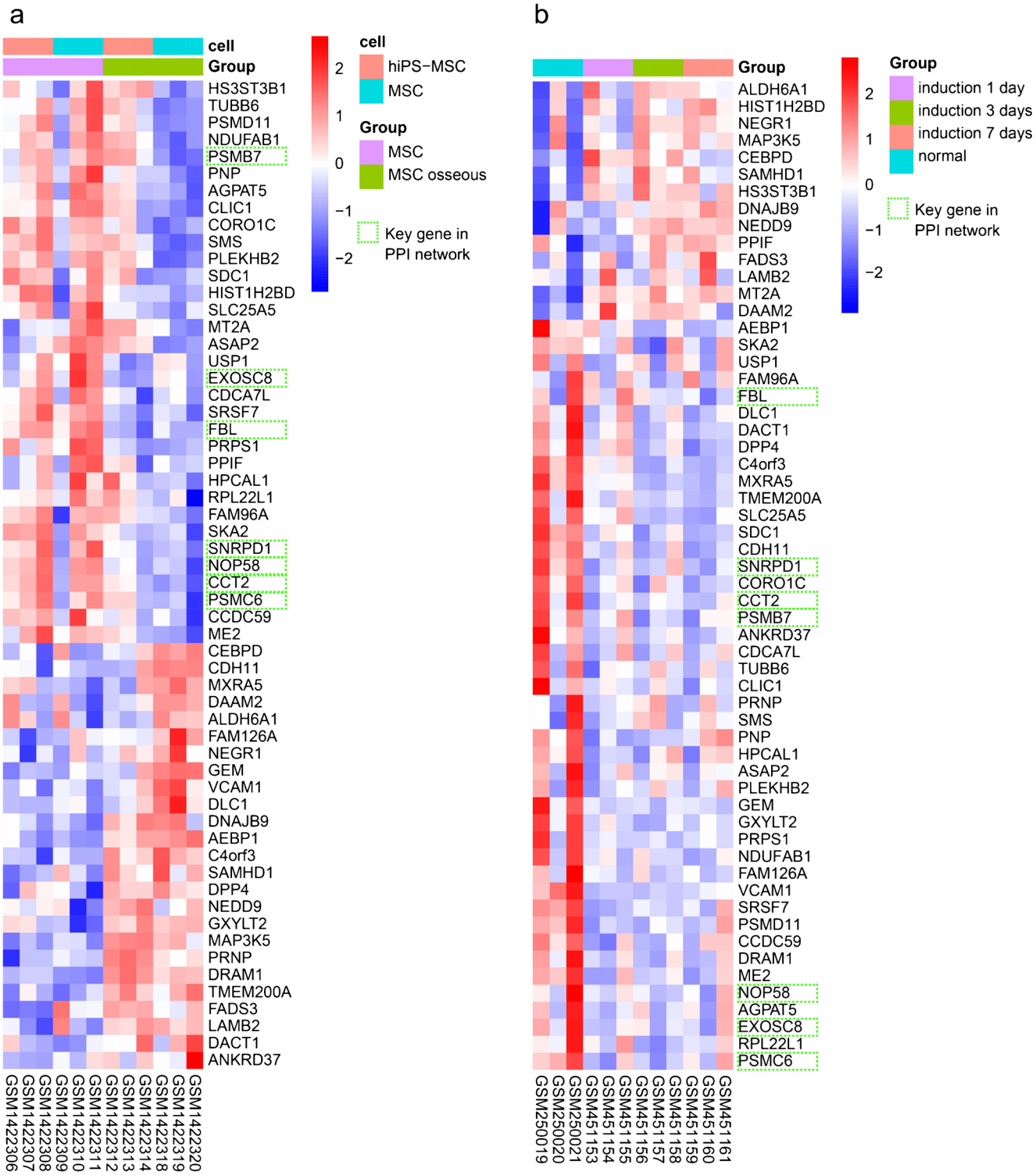


Fig. 4 Heatmap of key genes selected in GSE18043 and GSE58919. (**a**) Expression in the MSC osteogenic (green) and MSC (pale violet) groups, including hiPS-MSC (pink) and MSC (pale blue) in the two cell lines. Green border highlighting is the key gene for PPI screening. (**b**) Expression of 1 (pale violet), 3 (green) and 7 (orange) days and normal (pale blue) group by MSC chemically induced. Green border highlighting is the key gene for PPI screening

regulation of DNA, especially the centromeres, while there are different pathways involved in the KEGG pathway but the most significant is the regulation of the cell cycle and DNA. Pathway enrichments in GSE18043 however were used to

display the main pathway induced by osteo-inductive fluid. In the enrichment of the GSE18043 blue module, the top 10 GO terms were involved in the regulation of cell apoptosis and proliferation and RNA polymerase II promoter, while the top 10 KEGG pathways were focused on the regulatory pathways of cancer and stem cells (Fig. 5). These results showed that the dominant pathway of osteogenic differentiation induced by different inducements was different.

The key genes of the key modules related to different inducement modes were also enriched and analyzed. Enrichment of these key genes showed that the Wnt signaling pathway, ion response (including copper ion and calcium-mediated) and negative regulation of focal adhesion were involved in the top 10 GO terms while KEGG pathways were focused on proteasome and a variety of metabolisms (Fig. 6). Proteasome also appeared in GSE58919 brown module enrichment, which suggested that proteasome may play a key role in the regulation of osteogenesis.

PPI network construction

An interactive network of key genes involved in osteogenic differentiation of MSCs was successfully constructed. The most significant genes involved in the regulation of MSC osteogenic differentiation were found by screening out the hub node in the network. We found that CCT2, NOP58, FBL, PSMC6, PSMB7, EXOSC8 and SNRPD1 may be involved in MSC osteogenic differentiation as potential key genes (Fig. 7, Table 2).

Statistics of potential key genes selected from PPI network

We analyzed seven potential key genes from the PPI network, five of which differed significantly (CCT2, NOP58, FBL, EXPSC8 and SNRPD1) after 3D culture and osteogenic induction, especially FBL (Fig. 8).

Biological function of key gene

The results of GSEA indicated that changes in CCT2 had no meaningful pathway (Fig. 9a); changes in NOP58 were most likely to influence "KEGG_NOD_LIKE_RECEPTOR_SIGNALING_PATHWAY," "KEGG_ALZHEIMERS_DISEASE," "KEGG_HUNTINGTONS_DISEASE" and "KEGG_PROTEASOME" (Fig. 9b); changes in both FBL and EXOSC8 were most likely to cause changes in "KEGG_PYRIMIDINE_METABOLISM" and "KEGG_PURINE_METABOLISM" pathways (Fig. 9c,d); and changes in SNRPD1 were most likely to influence "KEGG_SPLICEOSOME" (Fig. 9e). SNRPD1 was a key molecule in "KEGG_SPLICEOSOME" pathway in GSEA analysis (Fig. 10a). Both PSMC6 and PSMB7 played a key role in "KEGG_PROTEASOME" pathway by enrichment of the NOP58 (Fig. 10b).

Discussion

According to previous studies of osteogenic differentiation, we found that two key factors are needed to be taken into account—the type of cell and cell-inducer mode. These factors are considered specific in the process of cell differentiation with a specific subtle influence. There are multiple paths in the regulation of MSC differentiation to find the intersection of different paths. Such an intersection is the key point for the foreseeable. So, we selected GSE58919 and GSE18043 in the GEO database. The former consisted of two cell lines, which were cultured in 3D culture with growth factors, while the latter was observed only on days 1, 3 and 7 and normally cultured in osteo-inductive fluid. WGCNA can connect character to the data, which is very helpful in exploring the above-mentioned problem. By taking the intersection of the key module genes of the two datasets, the key genes could induce osteogenic differentiation in different MSC cell lines, which could be screened out in different ways (Figs. 1, 2, 3).

To understand the function of these key module genes and the key genes screened out, GO and KEGG pathway enrichment analyses were carried out. After pathway enrichments of the key modules, we found that the main participation pathways of different inductive modes were different. Proteasome was still found in the pathway enrichments of the key genes, suggesting proteasome plays a key role in osteogenic differentiation. It is worth mentioning that the results of enrichment pathways of the key genes showed that adhesion and the Wnt signaling pathway played an important role in the regulation of osteogenic differentiation of cells (Gao et al. 2017). Liu et al. (2013) found that matrix culture and osteo-inductive fluid culture had a low extent but similar rhythm to induce osteogenic differentiation, which affirmed the key role of adhesion and the Wnt signaling pathway in osteogenic differentiation. There are also various ion-response pathways, especially the calcium-mediated pathway, which is the key pathway to regulate bone remodeling by the signal of calcium ion. Osteoblasts and osteoclasts have calcium-sensitive receptors that can sense the changes in intracellular and extracellular calcium concentrations. The study found that when the concentration of extracellular calcium is increased, it can increase the level of osteogenic differentiation markers and promote the differentiation of BMSCs into osteoblasts. These pathways cross-link each other to form a network. For example, the activation of the Wnt pathway requires the opening of calcium channels and the involvement of PKC/ERK. Channel proteins need to undergo thiogenation at specific cysteine sites in order to play a biological role (Liu et al.

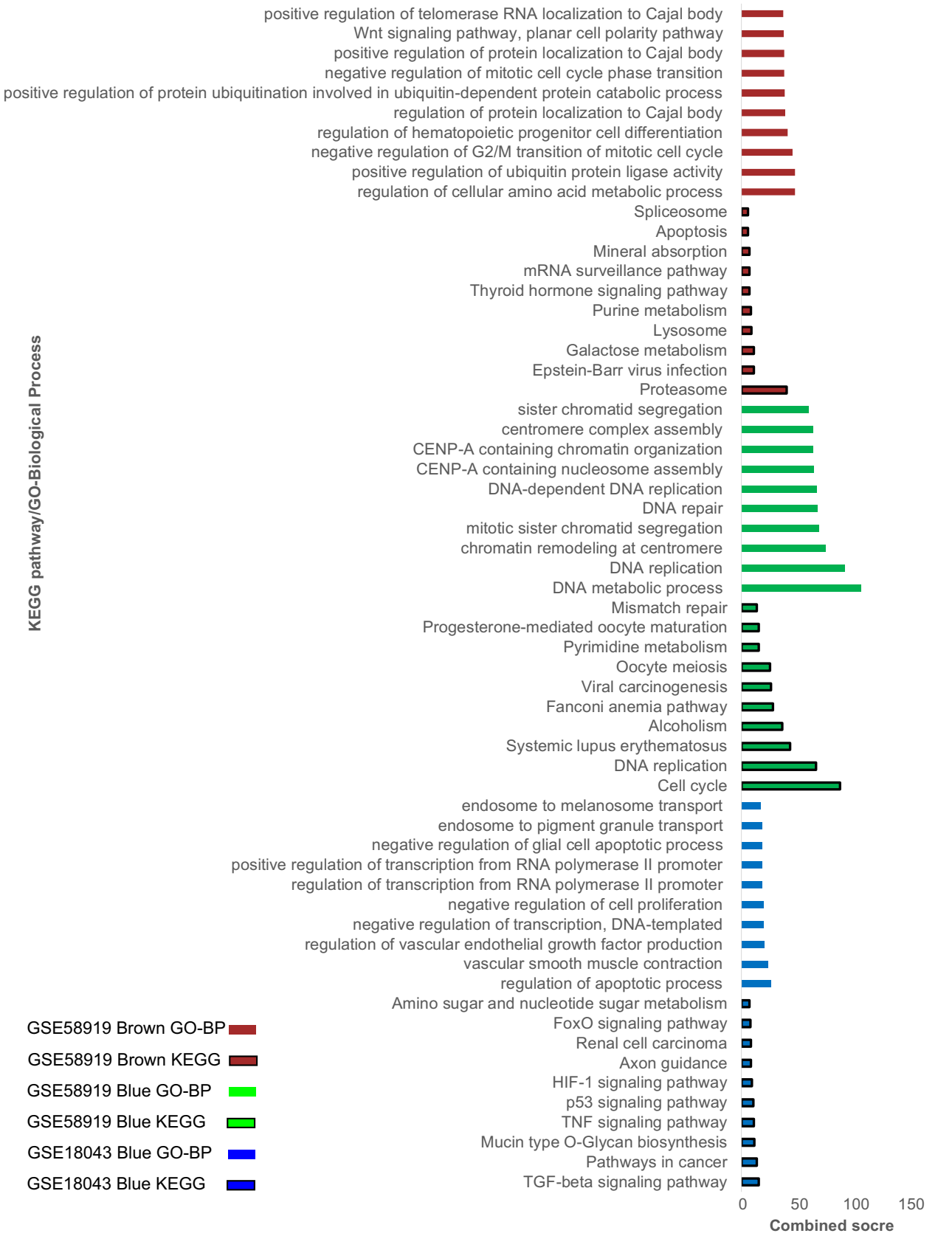
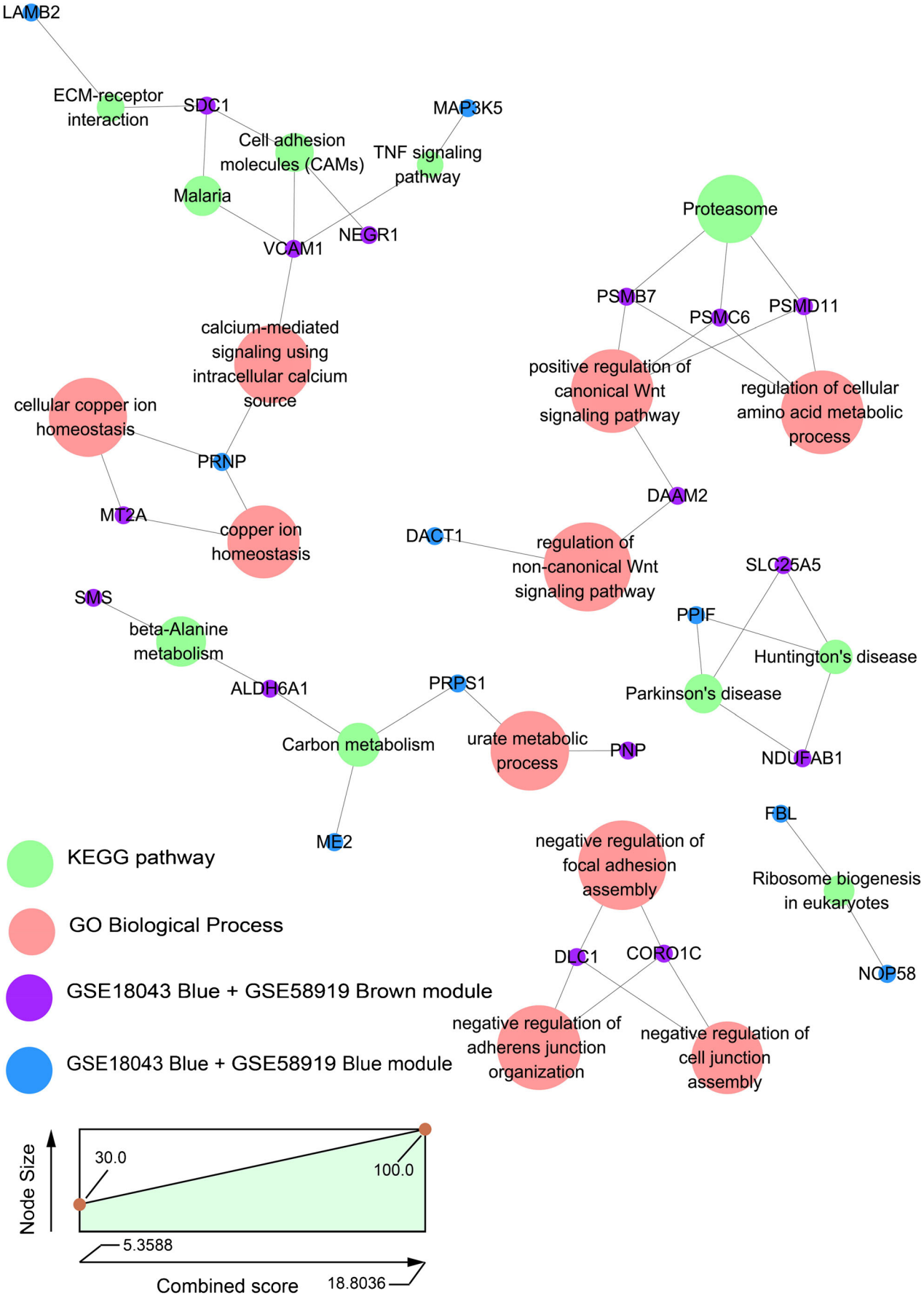


Fig. 5 GO and KEGG pathway enrichment analyses of module-specific genes. The column color (brown, green, blue) represents the respective modules (GSE58919 brown, GSE58919 blue, and GSE18043 blue); the frameless column represents the GO biological process; and the column with black border represents the KEGG pathway


LAMB2
ECM-receptor interaction
SDC1
MAP3K5
Cell adhesion molecules (CAMs)
TNF signaling pathway
Malaria
NEGR1
VCAM1
calcium-mediated signaling using intracellular calcium source
cellular copper ion homeostasis
PRNP
MT2A
copper ion homeostasis
Proteasome
PSMB7
PSMC6
PSMD11
positive regulation of canonical Wnt signaling pathway
regulation of cellular amino acid metabolic process
DAAM2
DACT1
regulation of non-canonical Wnt signaling pathway
SLC25A5
PPIF
Huntington's disease
Parkinson's disease
NDUFAB1
SMS
beta-Alanine metabolism
ALDH6A1
PRPS1
Carbon metabolism
urate metabolic process
PNP
ME2
FBL
Ribosome biogenesis in eukaryotes
NOP58
negative regulation of focal adhesion assembly
DLC1
CORO1C
negative regulation of adherens junction organization
negative regulation of cell junction assembly
KEGG pathway
GO Biological Process
GSE18043 Blue + GSE58919 Brown module
GSE18043 Blue + GSE58919 Blue module
Node Size
30.0
100.0
5.3588
Combined score
18.8036

◀ **Fig. 6** GO and KEGG pathway enrichment analyses of the key genes. Purple nodes represent GSE18043 blue and GSE58919 brown overlapping genes; light blue nodes represent GSE18043 blue and GSE58919 blue overlapping genes; blue nodes represent the KEGG pathway; red nodes represent the GO biological process. The size of blue and red nodes is determined by the combined score

2014). Endogenous hydrogen sulfide is an essential substance for thiolation. Estrogen increases intracellular hydrogen sulfide levels by up-regulating cystathionine-β-synthase (CBS) and cystathionine-γ lyase (CSE) levels, thus regulating the Wnt pathway.

Our finding of the key genes between different osteogenic induction and different MSC cell lines suggests that CCT2, NOP58, FBL, PSMC6, PSMB7, EXOSC8 and SNRPD1 might regulate a common path in these different situations to regulate MSC osteogenic differentiation. Furthermore, CCT2, NOP58, FBL, EXPSC8 and SNRPD1 were significantly different after osteogenic induction ($p < 0.05$). Understanding the specific function of genes in osteogenic regulation is necessary; so, we chose single-gene GSEA analysis to understand these genes in depth. The single-gene GSEA analysis sets the phenotypic file with the expression of the target gene. Based on this idea, the most likely function of the gene was obtained. The results showed that NOP58 promoted "PROTEASOME" pathway in osteogenic differentiation and that PSMC6 and PSMB7 were the main participants in this pathway. Moreover, SNRPD1 was showed to be involved in regulating the "SPLICEOSOME" pathway and as the major participant in this pathway. It is worth mentioning that the functional interactions of these key molecules also confirm the accuracy of the PPI network we constructed.

Some of the results from this study are consistent with those of the previous study. For example, T-complex protein 1 subunit beta (CCT2) can regulate cell differentiation or death by influencing the cell cycle (Minegishi et al. 2018). Another is the report of Alves et al. (2010) that states CCT2 might be involved in the regulation of osteoblast

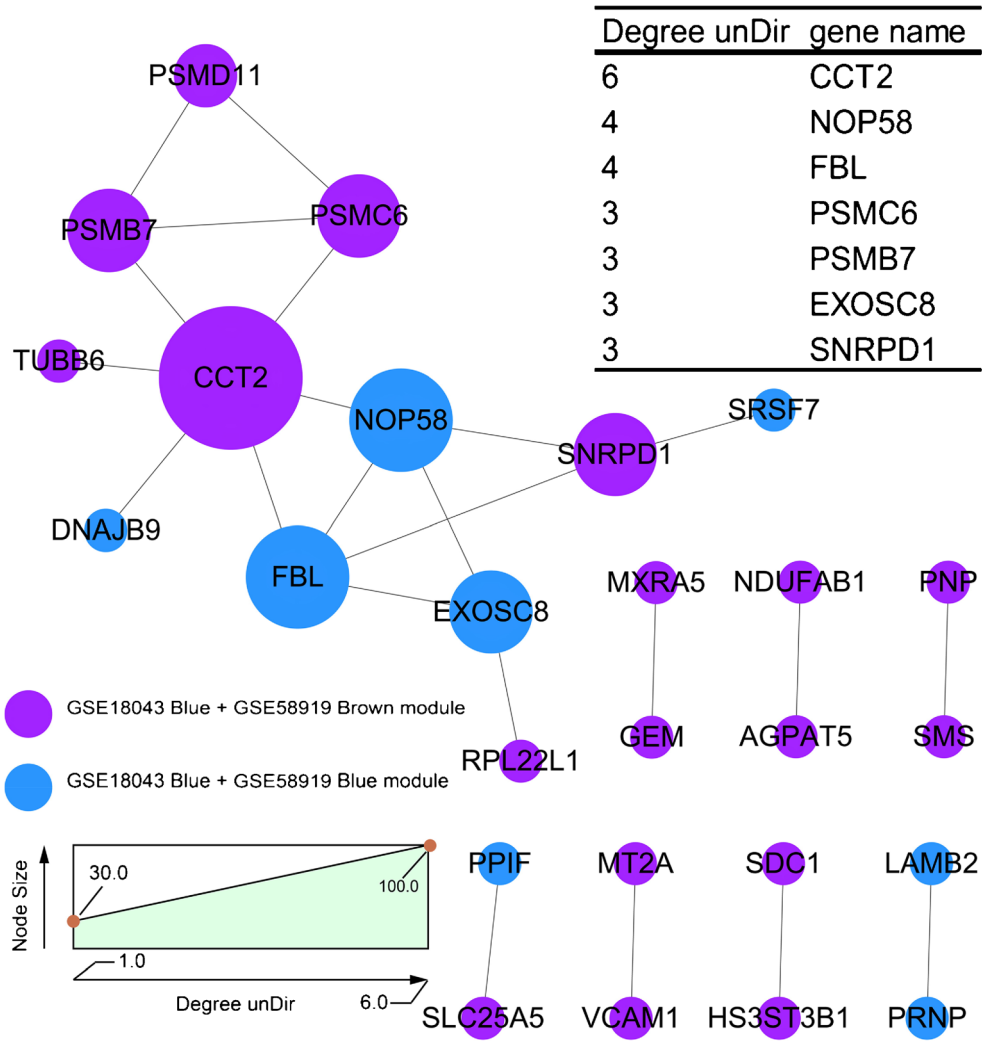


Fig. 7 PPI network of the key genes. Purple nodes represent GSE18043 blue and GSE58919 brown overlapping genes; light blue nodes represent GSE18043 blue and GSE58919 blue overlapping genes. The size of the nodes is determined by each node degree and it was displayed "Degree unDir" for the top 7 genes

Table 2 Characteristics of potential key genes in MSC osteogenic differentiation

Function	Reference
cell cycle and osteoblast differentiation	(Minegishi, et al., 2018, Alves, et al., 2010)
Influences 2'-O-methylation of rRNAs to regulate proliferation	(Qin, et al., 2017)
Influences 2'-O-methylation of rRNAs to regulate proliferation and differentiation	(Bouffard, et al., 2018, Watanabe-Susaki, et al., 2014, Shubina, et al., 2016)
Participates in stem cell differentiation and self-renewal	(Vilchez, et al., 2012, Koyuncu, et al., 2018, Saez, et al., 2018)
Corrects neuronal differentiation	(Giunta, et al., 2016, Lloret-Llinares, et al., 2018, Makino, et al., 2013)
interacts with cajal body to regulate transcription	(Smolinski and Kolowerzo, 2012)

mineralization or not. However, some of our results also differ from those in the previous study. In addition to CCT2, there was no direct evidence that the key genes were involved in the regulation of osteogenic differentiation. For example, fibrillarin (FBL) was down-regulated after osteogenic induction. It has been reported that fibrillarin (FBL) has the ability to regulate the differentiation (Bouffard et al. 2018; Watanabe-Susaki et al. 2014). The primary function of FBL is to participate in methylation and process pre-rRNA (Shubina et al. 2016). FBL and nucleolar protein 58 (NOP58) are core proteins of box C/D small nucleolar ribonucleoprotein complexes (snoRNPs) and they can influence 2′-O-methylation of rRNAs and regulate the proliferation and formation of tumor in vivo (Qin et al. 2017). Meanwhile, FBL and NOP58 are enriched in ribosome biogenesis in eukaryotes. In recent years, extensive literature has confirmed that ribosome biogenesis plays a key role in differentiation (Sanchez et al. 2016; Stedman et al. 2015), including osteoblast differentiation (Neben et al. 2017). Proteasome 26S subunit, ATPase 6 (PSMC6) and proteasome subunit beta 7 (PSMB7) are the essential subunits that contribute to completing the assembly of proteasome. PSMC6 and PSMB7 are involved in the Wnt signaling pathway, amino acid metabolic process and proteasome in enrichment analysis. The previous study suggested that proteasome could regulate the differentiation of human embryonic stem cells (hESCs) (Vilchez et al. 2012). With the development of stem cell research, the role of the ubiquitin–proteasome system was exposed (Koyuncu et al. 2018; Saez et al. 2018), which suggests that autophagy may be an important mechanism involved in the regulation of MSC differentiation. Exosome component 8 (EXOSC8) is essential for correcting neuronal differentiation (Giunta et al. 2016). It has been reported that exosome is involved in the regulation of stem cell differentiation (Lloret-Llinares et al. 2018) and that there is a strong correlation between exosome and proteasome in the regulation of the degradation of substance in cells (Makino et al. 2013). Small nuclear ribonucleoprotein D1 polypeptide (SNRPD1) interacts with the cajal body to regulate transcription (Smolinski and Kolowerzo 2012) (Table 2). Our results, however, did not show classical key genes in osteogenic differentiation, such as RUNX2, which were relatively novel molecules. This might be due to the analysis method we

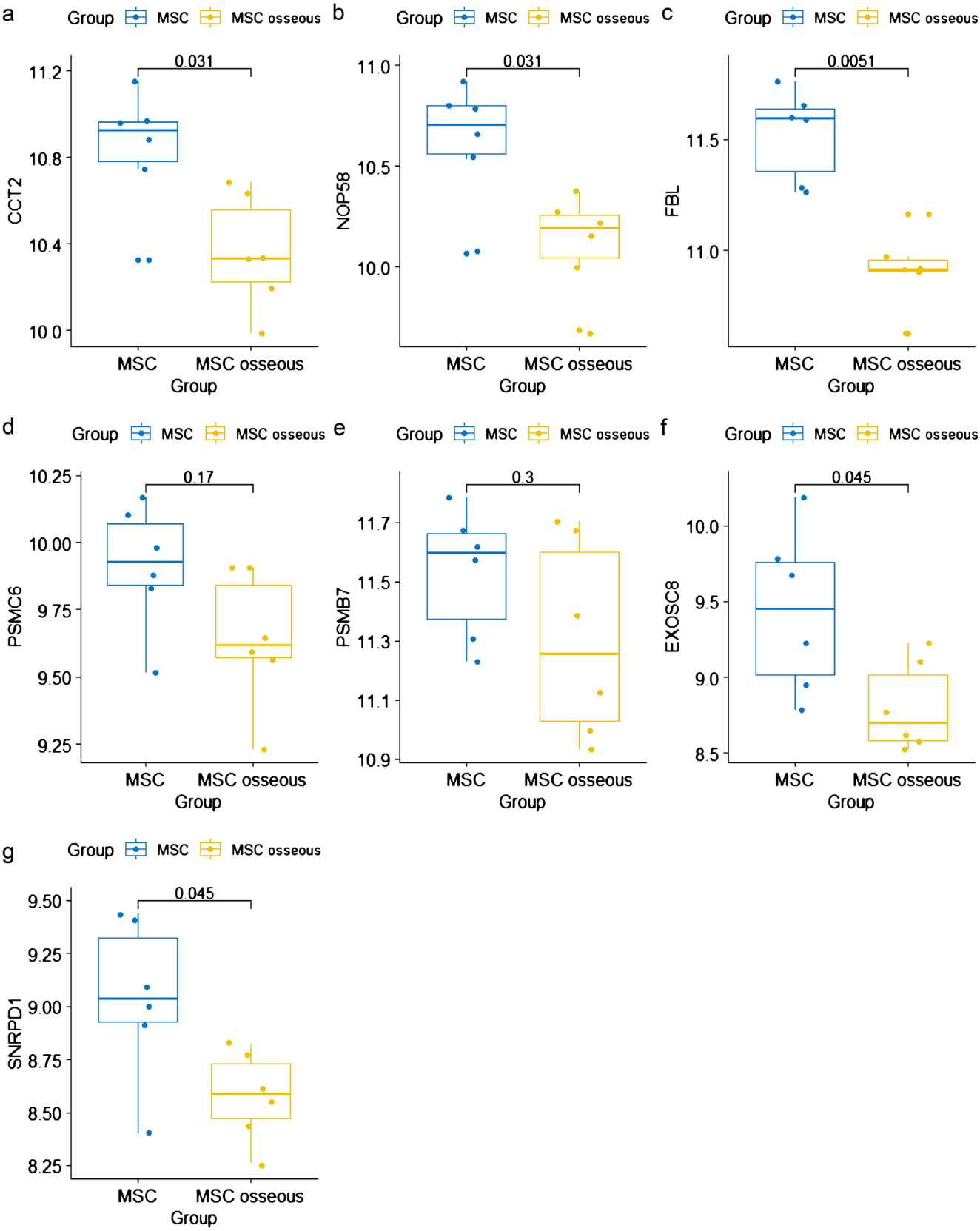


Fig. 8 Boxplots of seven key genes selected from the PPI network, including five genes with significant difference. (**a**) CCT2 ($p < 0.05$), (**b**) NOP58 ($p < 0.05$), (**c**) FBL ($p < 0.05$), (**d**) PSMC6 ($p > 0.05$), (**e**) PSMB7 ($p > 0.05$), (**f**) EXOSC8 ($p < 0.05$), (**g**) SNRPD1 ($p < 0.05$)

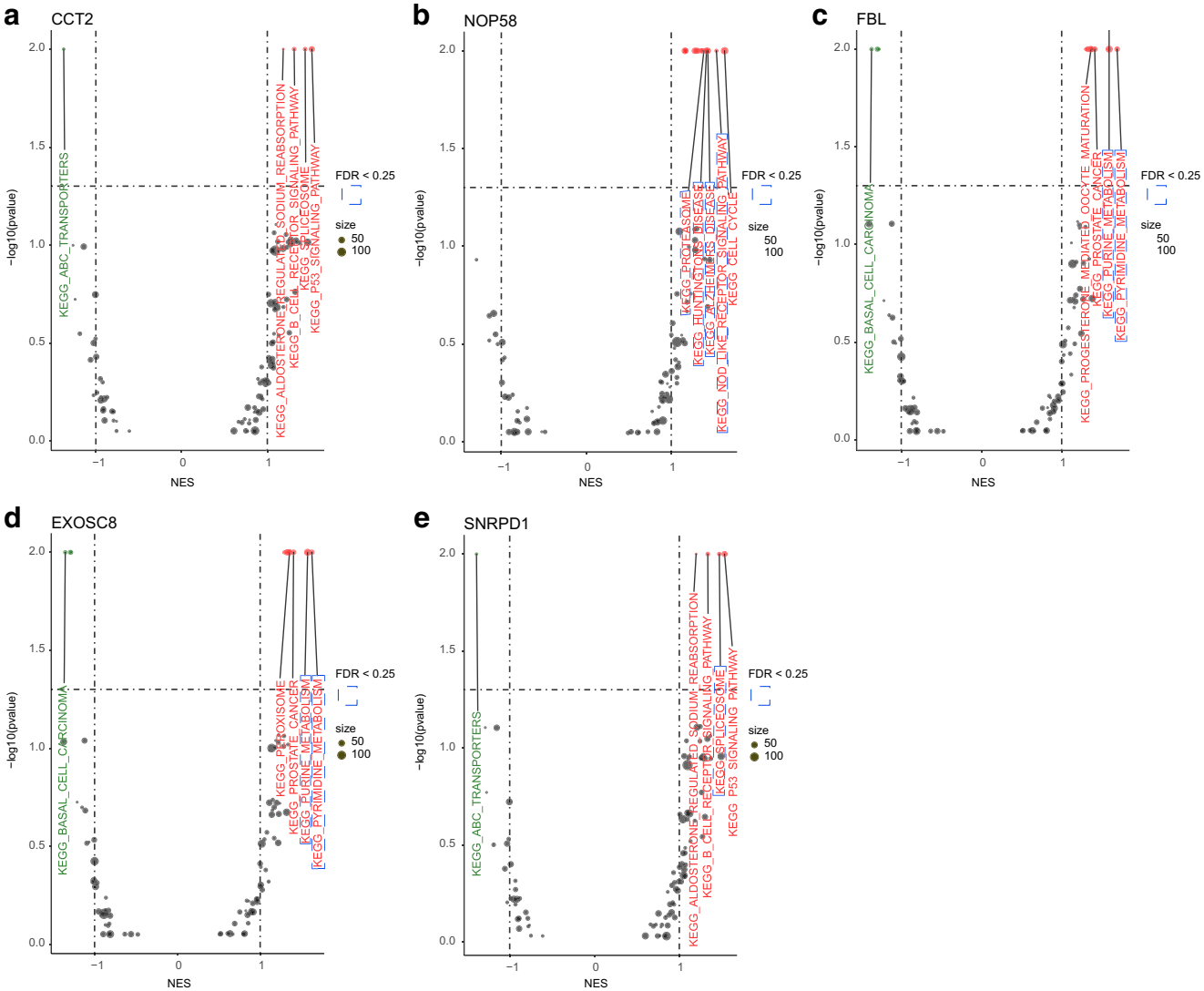


Fig. 9 Gene set enrichment analysis of promising targets. Biological function of (**a**) CCT2, (**b**) NOP58, (**c**) FBL, (**d**) EXOSC8 and (**e**) SNRPD1. The size of the dot was determined by the size of the enriched genes. The red dots indicated that upregulation of the function was significantly enriched in promising targets with the high-expression group. Dark green dots indicated that downregulation of the function was significantly enriched in promising targets with the low-expression group. NES, normalized enrichment score

selected; "WGCNA" cluster multiple genes according to the gene expression trend rather than gene expression difference. "WGCNA" could further identify key genes by connecting traits and gene clusters. Based on this idea, it may be possible to ignore some genes that express significantly but the expression pattern was not fit for the traits. The advantages of "WGCNA" based on gene expression patterns however will offer us a broader perspective into the mysteries.

In summary, we analyzed the RNA microarray data of MSC osteogenic differentiation under different conditions. Whether 3D scaffold with growth factors or biochemical agent-inducing osteogenic differentiation of MSCs is utilized, proteasome, adhesion, ion response, Wnt signaling pathway and ribosome can be used as the core to regulate osteogenic differentiation of MSCs. Seven potential key genes (CCT2, NOP58, FBL, PSMC6, PSMB7, EXOSC8 and SNRPD1) were found. In particular, FBL played a key role in MSC osteogenic differentiation. The key genes we found may be useful in understanding the common path in regulating osteogenic differentiation.

Acknowledgments We thank Kyrkou A and Murphy C for providing the GSE58919 data and Hamidouche Z, Fromigué O, Ringe J, Häupl T, Vaudin P, Srouji S, Livne E and Marie P for the GSE18043 data. We also thank the GEO database.

Funding information This study was funded by the National Natural Science Foundation of China (Grant No. 81572139).

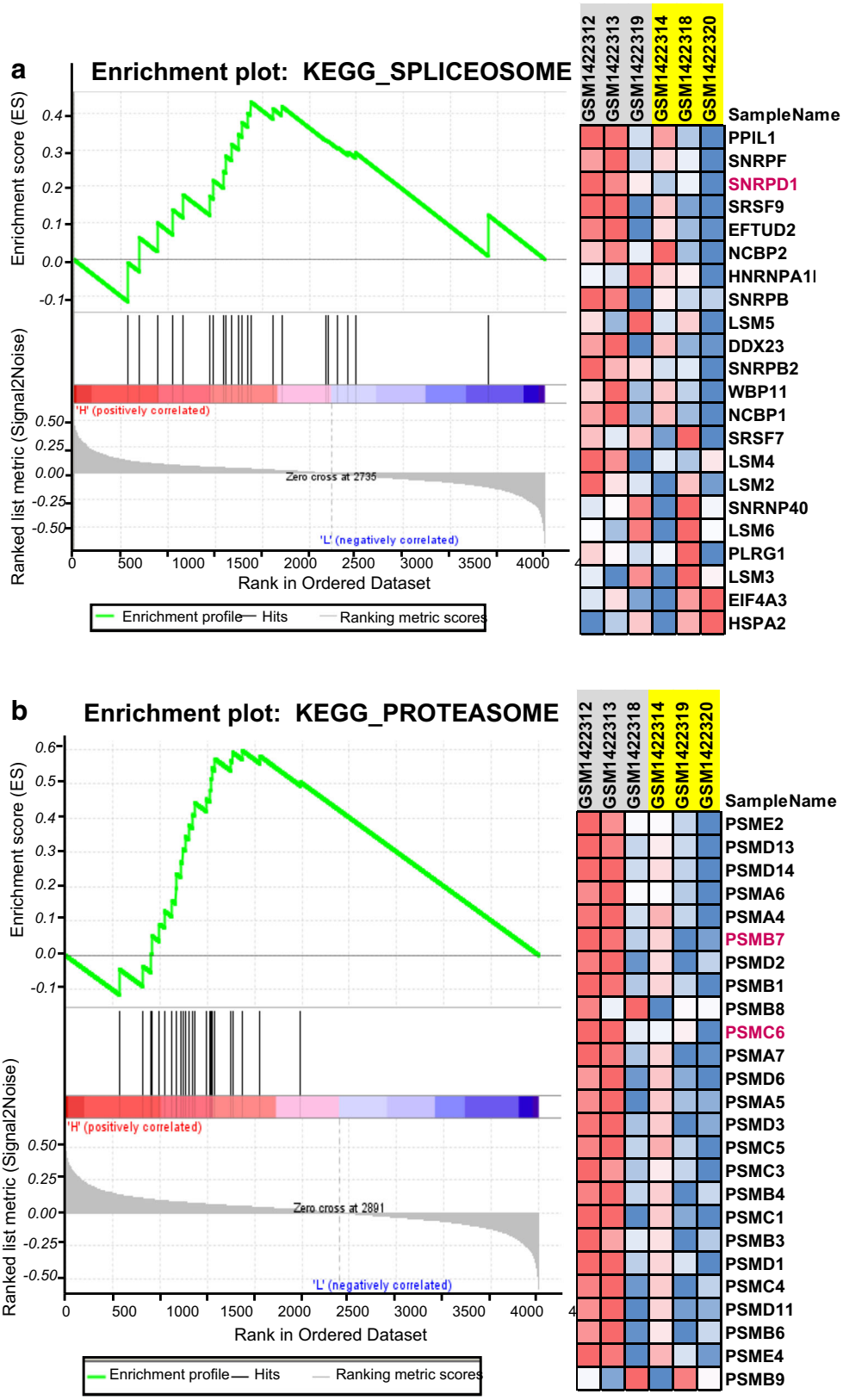


Fig. 10 Gene set enrichment analyses of SNRPD1 and NOP58. (**a**) GSEA shows that function of "SPLICEOSOME" was participated in SNRPD1 with high expression. (**b**) GSEA shows that function of "PROTEASOME" was participated in NOP58 with high expression. The expression level of the key genes involved in these pathways is shown. Red represents high expression; blue represents low expression. GSEA, gene set enrichment analysis

References

Alves RD, Eijken M, Swagemakers S, Chiba H, Titulaer MK, Burgers PC, Luider TM, van Leeuwen JP (2010) Proteomic analysis of human osteoblastic cells: relevant proteins and functional categories for differentiation. J Proteome Res 9:4688–4700

Bouffard S, Dambroise E, Brombin A, Lempereur S, Hatin I, Simion M, Corre R, Bourrat F, Joly JS, Jamen F (2018) Fibrillarin is essential for S-phase progression and neuronal differentiation in zebrafish dorsal midbrain and retina. Dev Biol 437:1–16

Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics 14:128

Chen X, Yan J, He F, Zhong D, Yang H, Pei M, Luo ZP (2018) Mechanical stretch induces antioxidant responses and osteogenic differentiation in human mesenchymal stem cells through activation of the AMPK-SIRT1 signaling pathway. Free Radic Biol Med 126: 187–201

Gao C, Peng S, Feng P, Shuai C (2017) Bone biomaterials and interactions with stem cells. Bone Res 5:17059

Giunta M, Edvardson S, Xu Y, Schuelke M, Gomez-Duran A, Boczonadi V, Elpeleg O, Muller JS, Horvath R (2016) Altered RNA metabolism due to a homozygous RBM7 mutation in a patient with spinal motor neuropathy. Hum Mol Genet 25:2985–2996

Golpanian S, Wolf A, Hatzistergos KE, Hare JM (2016) Rebuilding the damaged heart: mesenchymal stem cells, cell-based therapy, and engineered heart tissue. Physiol Rev 96:1127–1168

Hamidouche Z, Fromigue O, Ringe J, Haupl T, Vaudin P, Pages JC, Srouji S, Livne E, Marie PJ (2009) Priming integrin alpha5 promotes human mesenchymal stromal cell osteoblast differentiation and osteogenesis. Proc Natl Acad Sci U S A 106:18587–18591

Huebsch N, Lippens E, Lee K, Mehta M, Koshy ST, Darnell MC, Desai RM, Madl CM, Xu M, Zhao X, Chaudhuri O, Verbeke C, Kim WS, Alim K, Mammoto A, Ingber DE, Duda GN, Mooney DJ (2015) Matrix elasticity of void-forming hydrogels controls transplanted-stem-cell-mediated bone formation. Nat Mater 14:1269–1277

Killcoyne S, Carter GW, Smith J, Boyle J (2009) Cytoscape: a community-based framework for network modeling. Methods Mol Biol 563:219–239

Koyuncu S, Saez I, Lee HJ, Gutierrez-Garcia R, Pokrzywa W, Fatima A, Hoppe T, Vilchez D (2018) The ubiquitin ligase UBR5 suppresses proteostasis collapse in pluripotent stem cells from Huntington's disease patients. Nat Commun 9:2886

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, Ma'ayan A (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 44:W90–W97

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9:559

Li J, Huang Y, Song J, Li X, Zhang X, Zhou Z, Chen D, Ma PX, Peng W, Wang W, Zhou G (2018) Cartilage regeneration using arthroscopic flushing fluid-derived mesenchymal stem cells encapsulated in a one-step rapid cross-linked hydrogel. Acta Biomater

Liu W, Wei Y, Zhang X, Xu M, Yang X, Deng X (2013) Lower extent but similar rhythm of osteogenic behavior in hBMSCs cultured on nanofibrous scaffolds versus induced with osteogenic supplement. ACS Nano 7:6928–6938

Liu Y, Yang R, Liu X, Zhou Y, Qu C, Kikuiri T, Wang S, Zandi E, Du J, Ambudkar IS, Shi S (2014) Hydrogen sulfide maintains mesenchymal stem cell function and bone homeostasis via regulation of Ca(2+) channel sulfhydration. Cell Stem Cell 15:66–78

Lloret-Llinares M, Karadoulama E, Chen Y, Wojenski LA, Villafano GJ, Bornholdt J, Andersson R, Core L, Sandelin A, Jensen TH (2018) The RNA exosome contributes to gene expression regulation during stem cell differentiation. Nucleic Acids Res 46:11502–11513

Makino DL, Halbach F, Conti E (2013) The RNA exosome and proteasome: common principles of degradation control. Nat Rev Mol Cell Biol 14:654–660

Mathiasen AB, Qayyum AA, Jorgensen E, Helqvist S, Fischer-Nielsen A, Kofoed KF, Haack-Sorensen M, Ekblond A, Kastrup J (2015) Bone marrow-derived mesenchymal stromal cell treatment in patients with severe ischaemic heart failure: a randomized placebo-controlled trial (MSC-HF trial). Eur Heart J 36:1744–1753

McMillan A, Nguyen MK, Gonzalez-Fernandez T, Ge P, Yu X, Murphy WL, Kelly DJ, Alsberg E (2018) Dual non-viral gene delivery from microparticles within 3D high-density stem cell constructs for enhanced bone tissue engineering. Biomaterials 161:240–255

Minegishi Y, Nakaya N, Tomarev SI (2018) Mutation in the zebrafish cct2 gene leads to abnormalities of cell cycle and cell death in the retina: a model of CCT2-related leber congenital amaurosis. Invest Ophthalmol Vis Sci 59:995–1004

Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, Groop LC (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 34:267–273

Neben CL, Lay FD, Mao X, Tuzon CT, Merrill AE (2017) Ribosome biogenesis is dynamically regulated during osteoblast differentiation. Gene 612:29–35

Qin W, Lv P, Fan X, Quan B, Zhu Y, Qin K, Chen Y, Wang C, Chen X (2017) Quantitative time-resolved chemoproteomics reveals that stable O-GlcNAc regulates box C/D snoRNP biogenesis. Proc Natl Acad Sci U S A 114:E6749–e6758

Saez I, Koyuncu S, Gutierrez-Garcia R, Dieterich C, Vilchez D (2018) Insights into the ubiquitin-proteasome system of human embryonic stem cells. Sci Rep 8:4092

Sanchez CG, Teixeira FK, Czech B, Preall JB, Zamparini AL, Seifert JR, Malone CD, Hannon GJ, Lehmann R (2016) Regulation of ribosome biogenesis and protein synthesis controls germline stem cell differentiation. Cell Stem Cell 18:276–290

Sassoli C, Vallone L, Tani A, Chellini F, Nosi D, Zecchi-Orlandini S (2018) Combined use of bone marrow-derived mesenchymal stromal cells (BM-MSCs) and platelet rich plasma (PRP) stimulates proliferation and differentiation of myoblasts in vitro: new therapeutic perspectives for skeletal muscle repair/regeneration. Cell Tissue Res 372:549–570

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504

Shuai Y, Mao C, Yang M (2018) Protein nanofibril assemblies templated by graphene oxide nanosheets accelerate early cell adhesion and induce osteogenic differentiation of human mesenchymal stem cells. ACS Appl Mater Interfaces

Shubina MY, Musinova YR, Sheval EV (2016) Nucleolar methyltransferase fibrillarin: evolution of structure and functions. Biochemistry (Mosc) 81:941–950

Smolinski DJ, Kolowerzo A (2012) mRNA accumulation in the Cajal bodies of the diplotene larch microsporocyte. Chromosoma 121:37–48

Stedman A, Beck-Cormier S, Le Bouteiller M, Raveux A, Vandormael-Pournin S, Coqueran S, Lejour V, Jarzebowski L, Toledo F, Robine S, Cohen-Tannoudji M (2015) Ribosome biogenesis dysfunction leads to p53-mediated apoptosis and goblet cell differentiation of mouse intestinal stem/progenitor cells. Cell Death Differ 22:1865–1876

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102:15545–15550

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, von Mering C (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43:D447–D452

Vilchez D, Boyer L, Morantte I, Lutz M, Merkwirth C, Joyce D, Spencer B, Page L, Masliah E, Berggren WT, Gage FH, Dillin A (2012) Increased proteasome activity in human embryonic stem cells is regulated by PSMD11. Nature 489:304–308

Watanabe-Susaki K, Takada H, Enomoto K, Miwata K, Ishimine H, Intoh A, Ohtaka M, Nakanishi M, Sugino H, Asashima M, Kurisaki A (2014) Biosynthesis of ribosomal RNA in nucleoli regulates pluripotency and differentiation ability of pluripotent stem cells. Stem Cells 32:3099–3111

Yan L, Jiang B, Li E, Wang X, Ling Q, Zheng D, Park JW, Chen X, Cheung E, Du X, Li Y, Cheng G, He E, Xu RH (2018) Scalable generation of mesenchymal stem cells from human embryonic stem cells in 3D. Int J Biol Sci 14:1196–1210