Identifying candidate genes involved in osteoarthritis through bioinformatics analysis

X. Zhang, Z. Yuan, S. Cui

Department of Microsurgery, The First Affiliated Hospital of the Henan University of Science and Technology, Luoyang, Henan Province, China.

Abstract

Objective

This study aims to identify candidate genes and critical pathways involved in osteoarthritis (OA).

Methods

Gene expression data of synovial membrane from OA patients and normal controls (NCs) were downloaded from database. Totally, 15 OA and 14 NC chips were available. Differentially expressed genes (DEGs) were identified through limma package (log_2 fold change >0.585, false discovery rate (FDR) < 0.05), and protein-protein interaction (PPI) network was constructed using STRING. Moreover, perturbation and pathway enrichment analyses were performed through PerturbationAnalyzer in Cytoscape (iterative criteria $<1\times e^{-10}$) and clusterProfiler package (FDR <0.05), respectively.

Results

Totally, 236 up-regulated and 290 down-regulated DEGs were identified. In PPI network, 10 hub genes were found, including VEGFA, IL6, JUN, IL1B, ICAM1, ATF3, IL8, EGR1, CDKN1A, and JUNB. After perturbation analysis, 32 DEGs were passively and significantly changed, like PISD, RARRES3, EIF4G1, and EPHA3. Furthermore, 526 DEGs were enriched in 176 pathways, and pathway cross-talk network was constructed, involving 12 pathways and 66 cross-talks.

Conclusion

Pathways like rheumatoid arthritis, osteoclast differentiation, and cytokine-cytokine receptor interaction might play critical roles in OA, and previously unreported genes VEGFA, JUN, JUNB, PISD, RARRES3, EIF4G1, and EPHA3 might participate in OA, providing novel directions for drug targeting.

Key words

osteoarthritis, differentially expressed genes, protein-protein interaction, perturbation analysis, pathway cross-talk

Xinying Zhang, MM,
Zhengjiang Yuan, BM
Shuo Cui, MM

Please address correspondence to:
Xinying Zhang,
Microsurgical Ward Section,
The First Affiliated Hospital of the Henan University of Science and Technology,
NO.24 Jinghua Road,
Luoyang 471000,
Henan Province, China
E-mail: zhangxinying01@163.com

Received on April 15, 2015; accepted in revised form on November 23, 2015.

Competing interests: none declared.

Introduction

Osteoarthritis (OA) is a common form of chronically degenerative joint disease that causes loss of gross cartilage, inflammation of synovial membrane, remodeling of sub-articular bone, and formation of osteophytes (1). It affects about 27 million people in the United States and nearly 8 million people in the United Kingdom (2). OA may happen in fingers, feet, spine, hips, and knees, and the joint and periarticular abnormalities in OA can be evaluated by using ultrasound imaging (3). OA patients suffer from joint pain, stiffness, and effusion, which finally lead to muscle spasms, tendon contractions, and ligaments laxity. Typically, high humidity, cold temperature, and drop in barometric pressure cause increased pain.

Triggering factors of OA are obesity (4), mechanical injury, sex hormone levels (5), impairment of peripheral nerves (6), and genetic factors (7). It's reported that the C allele of rs3815148 on chromosome 7q22 was significantly associated with increased prevalence of knee and hand OA (8). In OA, the cells in synovial membrane are activated, producing chemokines, inflammatory cytokines, and other inflammatory mediators (9). Pro-inflammatory cytokines, including interleukin 6 (IL6), IL8, IL15, IL17, IL18, IL21, leukaemia inhibitory factor, IL1β, and tumour necrosis factor (TNF) have been shown to be implicated in OA pathophysiology (10). Among these genes, IL1β and TNF control the degeneration of articular cartilage matrix. Especially, IL8 is a pro-inflammatory mediator that presents in the synovial fluid of OA patients, and it correlates with inflammatory markers like IL6 and TNF-α (11). In addition, the aberrant expression of aggrecanases, collagenases, cyclooxygenase 2 (COX2), nitric oxide synthase 2 (NOS2), cysteine cathepsins B and S, serine proteinases, discoidin domain receptor 2 (DDR2), matrix metalloproteinase 1 (MMP1), MMP3, and MMP13 is involved in OA progression (12). DDR2 and MMP13 are associated with cleavage and degradation of type II collagen, while serine proteinases like HTRA1 are responsible for exposing DDR2 to the activation by type II collagen through degrading the pericellular matrix components (13). Moreover, chondrocytes in OA cartilage express IL1, caspase 1, and type 1 IL1 receptor (IL1RI). IL1 induces the expression of MMPs and participates in matrix depletion in OA cartilage, together with TNF-α, MMP1, MMP3, MMP8, and MMP13 (12). Additionally, platelet-derived growth factor (PDGF), bone morphogenetic proteins (BMP), and nerve growth factor (NGF) are involved in the process of pain (14). However, the etiology and pathogenesis underlying OA have not been elucidated clearly. Further work on identifying candidate genes and critical pathways involved in the initiation and progression of human OA might inform us with novel directions for targeted therapies.

In the present study, microarray meta-analysis was utilised to merge gene expression profiling data from two datasets. Differentially expressed genes (DEGs) between synovial membranes of OA patients and normal controls (NCs) were screened out, enrichment analysis was performed, protein-protein interaction (PPI) network was constructed, and network perturbation analysis and pathway cross-talk analysis were conducted.

Materials and methods

Microarray data

The raw gene expression data were acquired from GEO (Gene Expression Omnibus) repository (available at http://www.ncbi.nlm.nih.gov/projects/geo/) (15). In this study, only chips of OA patients and NCs in two datasets, GSE1919 (16) and GSE12021 (17), were used for bioinformatics analysis. What's more, the expression data in GSE12021 were detected through platform GPL96 and platform GPL97. However, platform GPL97 involves little gene symbols overlapped with other platforms, bringing difficulties to microarray meta-analysis. Therefore, only the chips detected through platform GPL96 were utilised in the present study. Totally, 29 chips were available, including 15 chips of OA patients and 14 chips of NCs (Table I).

Data preprocessing

Firstly, probe symbols were converted

Table I. The distribution of 29 chips.

Dataset	Platform	NC	OA
GSE1919	GPL91 [HG_U95A] Affymetrix Human Genome U95A Array	5	5
GSE12021	GPL96 [HG-U133A/B] Affymetrix Human Genome U133A/B Array	9	10

NC: normal control; OA: osteoarthritis.

to gene symbols, according to the annotation information in corresponding platforms. For gene symbols which were corresponding to multiple probes, expression values at probe-level were averaged to get gene expression values. Secondly, gene expression values were $\log_2$ transformed. Thirdly, missing values were estimated by using nearest neighbour averaging method in impute package (v. 1.40.0, available at http://www.bioconductor.org/packages/3.0/bioc/html/impute.html) (18). Fourthly, quantile method in preprocessCore package (v. 1.28.0, available at http://www.bioconductor.org/packages/3.0/bioc/html/preprocessCore.html) (19) was utilised to perform normalisation between chips. Finally, expression data in two datasets were merged by using CONOR package from the Bioconductor project (20).

DEGs screening and hierarchical clustering analysis

Genes differentially expressed between OA and NC samples were identified by utilising limma package in R language (v. 3.22.7, available at http://www.bioconductor.org/packages/3.0/bioc/html/limma.html) (21), and *p*-value was adjusted through Benjamini-Hochberg (BH) method (22), producing false discovery rate (FDR). The criteria for this analysis were $|\log_2$ fold change (FC$|$) >0.585 and FDR <0.05. Then, pheatmap package in R language was utilised to perform bidirectional hierarchical clustering analysis (BHCA) (23, 24). Based on Euclidean distances of expression values, DEGs with similar expression pattern were clustered.

The construction of PPI network

The PPIs among DEGs were investigated through online sever STRING (Search Tool for the Retrieval of Interacting Genes, available at http://string-db.org/) (25), and PPIs with combined score > 0.4 were up-loaded to Cytoscape (v. 2.8, available at http://cytoscape.org/) (26) to construct PPI network.

Perturbation analysis of PPI network

Analysis of the dynamic changes in PPI network caused by single- or multi-perturbation might provide directions for the discovery of key genes in a network. The perturbation analysis of PPI network was conducted by using PerturbationAnalyzer plugin (27) in Cytoscape software. The pattern for this analysis was multi-perturbation, the dissociation constant was set as 10, and the iterative criteria was $<1\times e^{-10}$.

Enrichment and pathway cross-talk analyses

To investigate the bio-functions that involve the identified DEGs, GO (gene ontology) functional enrichment analysis and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analysis were performed by using clusterProfiler package (v. 2.0.1, available at http://www.bioconductor.org/packages/3.0/bioc/html/clusterProfiler.html) in R language (28). The criterion for this analysis was FDR <0.05. Pathway cross-talks among enriched pathways were analysed by performing latent pathway identification analysis (LPIA) (29). Firstly, the enriched GO biological process (BP) terms and KEGG pathways were abstracted. Secondly, the bipartite network of GO BP terms and KEGG pathways was constructed through Cytoscape software, and edge in this network represented that genes involved in a certain pathway were also involved in the connected GO BP term. Edge weight of bipartite network was calculated by using the following formula:

$$w_{GP} = \left|\frac{G \cap P}{G \cup P}\right| \times med\{DE_x : x \in G \cap P\}$$

In this formula, P stands for KEGG pathway, G represents GO BP term, $\left|\frac{G \cap P}{G \cup P}\right|$ is the Jaccard similarity coefficient between P and G, DE denotes the expression value of DEG, $G \cup P$ signifies the DEGs involved in P or G, and $G \cap P$ shows the DEGs involved in both P and G. Thirdly, pathway cross-talk network was constructed by linking the pathways with at least one jointly connected GO BP term. In pathway cross-talk network, edge weight was calculated by utilising the following formula:

$$A_{ij} = \sum_{k=1}^{G} w_{G_k P_i} \times w_{G_k P_j}$$

Fourthly, the importance of each cross-talk between pathways was calculated by utilising random method:

$$T_{ij} = \frac{A_{ij}}{\sum_{j=1}^{N_p} A_{ij}}$$

In this formula, T_{ij} represents the possibility of linking pathway P_i to pathway P_j directly, and N_p is the count of pathways in this network. Furthermore, the *p*-value of pathway cross-talk was calculated by using bootstrap method (repeated sampling, from the first step to the fourth step) (30). The criterion for this analysis was *p*-value <0.05.

Results

DEGs screening and hierarchical clustering analysis

After data preprocessing and merging, the expression values of 8558 genes in 29 chips from 2 datasets were obtained. The expression pattern of 29 chips is shown in Figure 1A. The median values of these chips were aligned, indicating that the expression values from different datasets were well normalised and could be utilised for further analysis. After DEGs screening, a total of 526 DEGs ($|\log_2$ FC$|$ >0.585 and FDR <0.05) were identified between OA and NC samples, including 236 up-regulated and 290 down-regulated DEGs (Fig. 1B). After BHCA, DEGs with similar expression pattern were clustered, distinguishing OA from NC samples clearly (Fig. 1C).

The construction of PPI network

According to the PPIs among DEGs,

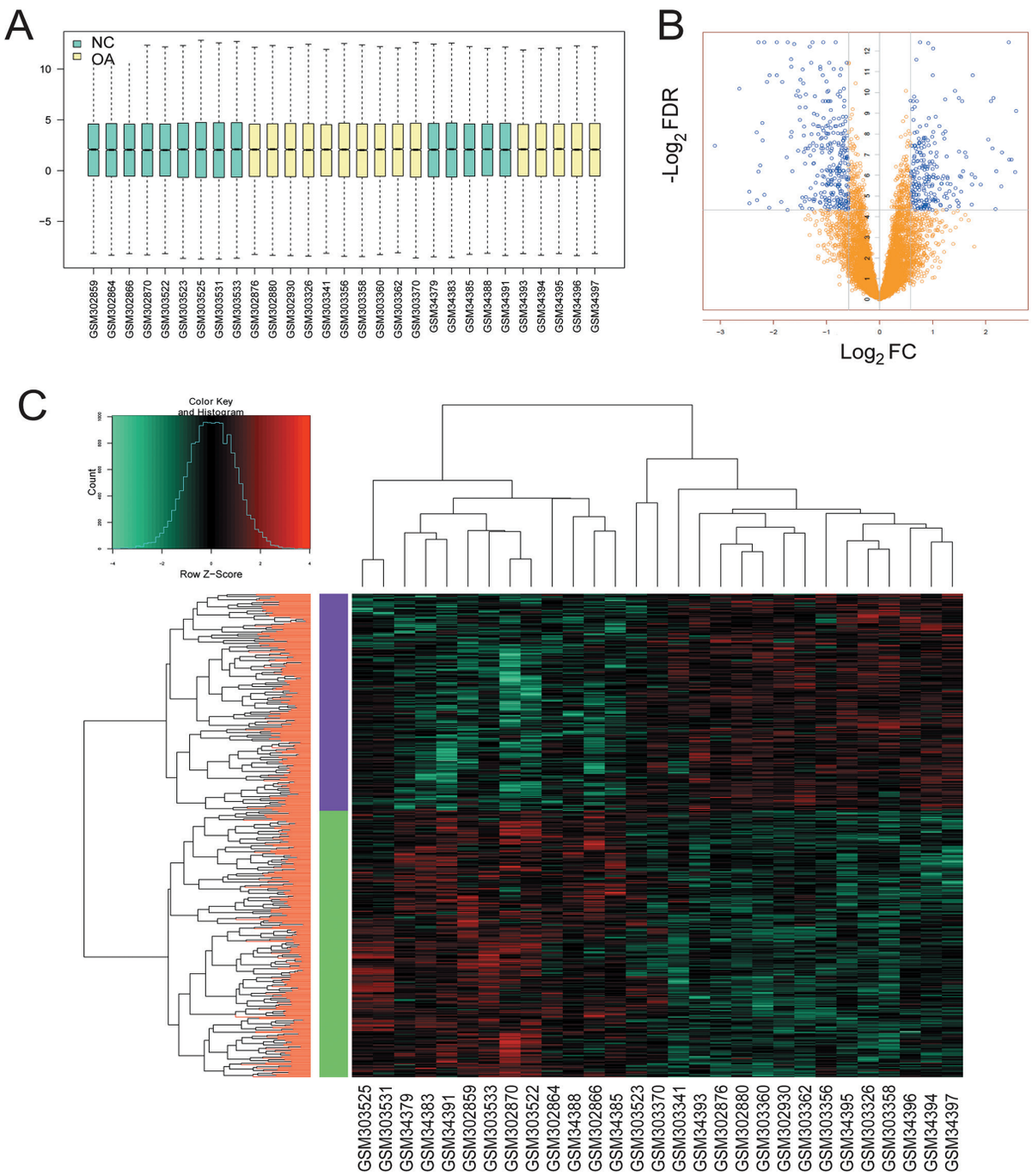

Fig. 1. The results of DEGs screening and hierarchical clustering analysis.
A. Gene expression data after merging and normalisation. **B.** The volcano plot of gene expression. **C.** Bidirectional hierarchical clustering analysis. DEGs: differentially expressed genes; NC: normal control; OA: osteoarthritis; FDR: false discovery rate; FC: fold change; the green to red bar in Figure 1C stands for the change from low expression to high expression.

PPI network was constructed (Fig. 2), involving 365 nodes (DEGs) and 1227 edges (PPIs, combined score >0.4). Among the DEGs in PPI network, 210 DEGs were down-regulated and 155 DEGs were up-regulated. The top 10 DEGs with most degree, which are defined as hub genes in PPI network, are listed in Table II.

Perturbation analysis of PPI network

The hub genes in PPI network were utilised to perform perturbation analysis (multi-perturbation, dissociation constant = 10, and iterative criteria $<1\times e^{-10}$), obtaining the perturbation sub-network (Fig. 3). The expression values of 32 DEGs were passively and significantly affected by perturbation (12 DEGs were down-regulated and 20 DEGs were up-regulated), while the expression values of 10 hub genes were elevated after perturbation. Among the passively changed DEGs, *PISD*, *RARRES3*, *EIF4G1*, and *EPHA3* significantly changed their expression values (change ratio >2). However, these 4 DEGs have no direct PPI relationship with the 10 hub genes.

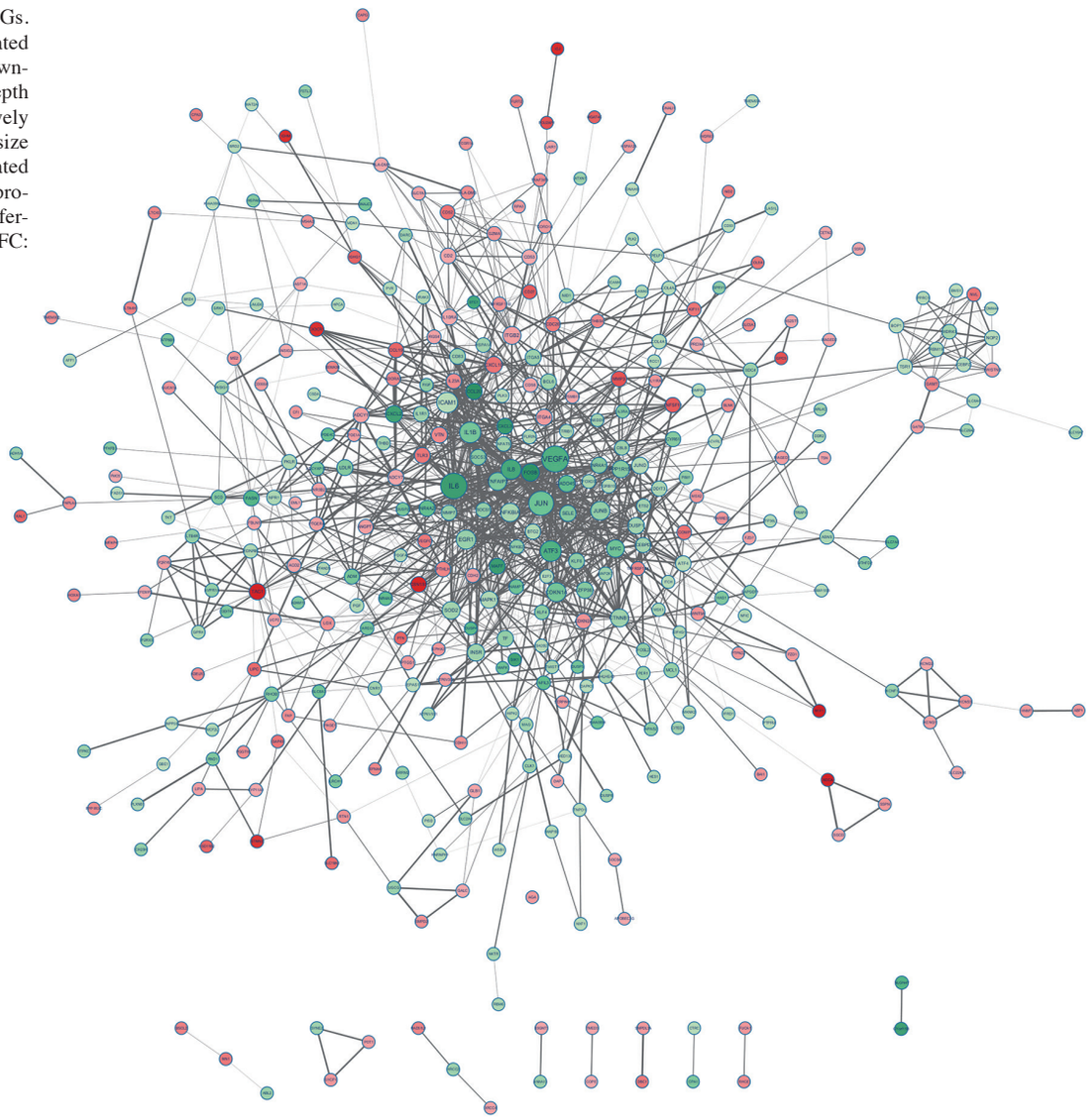

Fig. 2. PPI network of DEGs. Red nodes: up-regulated DEGs; green nodes: down-regulated DEGs; the depth of node colour was positively related with $|\log_2 FC|$; the size of node was positively related with degree; PPI: protein-protein interaction; DEGs: differentially expressed genes; FC: fold change.

Table II. The top 10 DEGs with most degree (hub genes).

Gene	Degree	Log_2FC	FDR	State
VEGFA	62	-1.93761416	0.000183029	Down-regulated
IL6	60	-2.282816391	0.013912616	Down-regulated
JUN	54	-1.535360267	0.000923929	Down-regulated
IL1B	38	-1.423515398	0.007149227	Down-regulated
ICAM1	37	-0.6456171	0.049335085	Down-regulated
ATF3	37	-1.935650407	0.000550951	Down-regulated
IL8	33	-2.080069775	0.046203768	Down-regulated
EGR1	31	-1.027871683	0.033167196	Down-regulated
CDKN1A	30	-1.436349242	0.000713816	Down-regulated
JUNB	29	-1.085945071	0.013431938	Down-regulated

DEGs: differentially expressed genes; FC: fold change; FDR: false discovery rate.

Enrichment and pathway cross-talk analyses

After enrichment analysis, up-regulated DEGs were significantly (FDR <0.05) enriched in functions about leukocyte migration, extracellular matrix, binding, and lipase, and pathways about rheumatoid arthritis, lysosome, phagosome, and cytokine-cytokine receptor interaction (Fig. 4). While, down-regulated DEGs were significantly (FDR <0.05) enriched in functions about developmental process, nucleus, and nucleic acid binding, and pathways about MAPK signaling, osteoclast differentiation, cancer, and malaria (Fig. 4). Totally, 526 DEGs were enriched in 176 pathways and 4973 GO BP terms, and the bipartite network of GO BP terms and KEGG pathways was constructed, involving 83012 edges. After linking the pathways with at least one jointly connected GO BP term, pathway cross-talk network was constructed, involving 12 pathways (p-value <0.05, Table III) and 66 cross-talks (Fig. 5). Among these pathways, rheumatoid arthritis, osteoclast differentiation, and cytokine-cytokine receptor interaction were enriched by plenty of DEGs, and had strong cross-talks with other pathways.

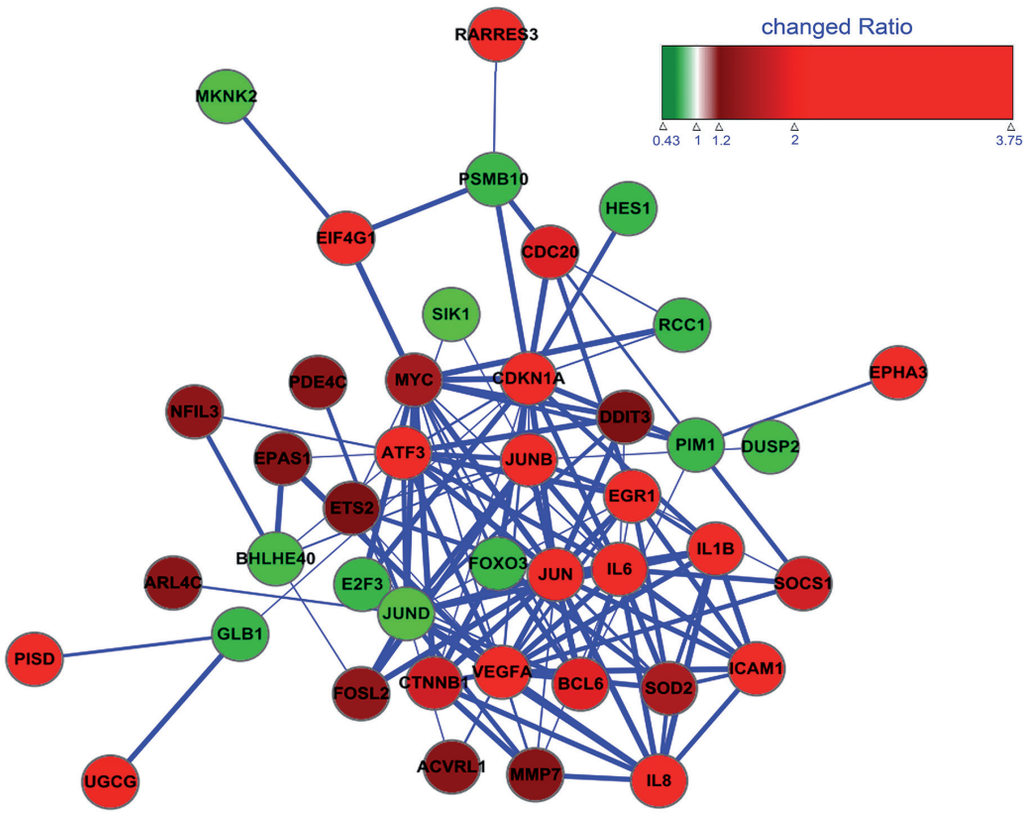

Fig. 3. The perturbation subnetwork of PPI network. Red nodes: up-regulated DEGs after perturbation; green nodes: down-regulated DEGs after perturbation; the depth of node colour was positively related with the ratio of expression value after perturbation to expression value before perturbation; PPI: protein-protein interaction; DEGs: differentially expressed genes.

Table III. The 12 pathways with cross-talk.

Pathway ID	Pathway name	*p*-value	DEGs
hsa05323	Rheumatoid arthritis	0	*ICAM1*, *PGF*, *ATP6V1B1*, *IL1B*, *JUN*, *IL8*, *IL6*, *CCL20*, *MMP3*, *TNFSF11*, *TNFRSF11A*, *HLA-DMB*, *IL23A*, *ATP6V0E2*, *CTSK*, *ANGPT1*, *ITGB2*
hsa05219	Bladder cancer	0.012	*DAPK3*, *E2F3*, *PGF*, *FIGF*, *CDKN1A*, *MYC*, *IL8*, *VEGFC*, *CDKN2A*
hsa05160	Hepatitis C	0.036	*MAPK11*, *NFKBIA*, *SOCS3*, *LDLR*, *CDKN1A*, *IL8*, *TLR3*
hsa05144	Malaria	0	*ICAM1*, *DARC*, *SDC4*, *HBD*, *IL1B*, *IL8*, *IL6*, *COMP*, *THBS4*, *ITGB2*
hsa04976	Bile secretion	0.047	*LDLR*, *SLCO1B1*, *ADCY1*, *NR0B2*, *ADCY7*
hsa04940	Type I diabetes mellitus	0.048	*IL1B*, *HLA-DMB*
hsa04380	Osteoclast differentiation	0	*JUND*, *MAPK11*, *NFKBIA*, *IL1R1*, *MAP2K7*, *FOSL2*, *SOCS1*, *NFKB2*, *JUNB*, *SOCS3*, *IL1B*, *JUN*, *FOSB*, *TNFSF11*, *TNFRSF11A*, *BLNK*, *TNFRSF11B*, *CTSK*
hsa04110	Cell cycle	0.03	*E2F3*, *CDKN1A*, *MYC*, *GADD45B*, *CDC20*, *CDKN2A*
hsa04060	Cytokine-cytokine receptor interaction	0.011	*BMPR2*, *PDGFA*, *IL1R1*, *FIGF*, *IL1B*, *INHBB*, *IL8*, *CXCL2*, *IL6*, *CXCL3*, *CCL20*, *CX3CR1*, *TNFSF11*, *CD27*, *CCL19*, *CXCL10*, *VEGFC*, *TNFRSF11A*, *IL23A*, *TNFRSF11B*, *IL11RA*, *IL10RA*
hsa03440	Homologous recombination	0.015	*XRCC2*, *TOP3B*, *RPA3*
hsa00600	Sphingolipid metabolism	0.033	*UGCG*, *SMPD2*, *GLB1*, *GALC*
hsa00330	Arginine and proline metabolism	0.044	*GAMT*, *ABP1*, *GATM*, *P4HA2*

ID: identifier; DEGs: differentially expressed genes.

Discussion

OA is a prevalent chronic joint disease. Patients with OA suffer from joint pain, stiffness, and effusion. In this study, microarray meta-analysis was performed to identify candidate genes and critical pathways involved in OA progression. Consequently, 526 DEGs were identified between OA and NC samples. Then PPI network was constructed, and 10 hub genes were identified, including *VEGFA*, *IL6*, *JUN*, *IL1B*, *ICAM1*, *ATF3*, *IL8*, *EGR1*, *CDKN1A*, and *JUNB*. Furthermore, perturbation analysis of PPI network was performed, and *PISD*, *RARRES3*, *EIF4G1*, and *EPHA3* were significantly changed by the perturbation of hub genes. Moreover, enrichment and pathway cross-talk analyses were conducted, and rheumatoid arthritis, osteoclast differentiation, and cytokine-cytokine receptor interaction had strong cross-talks with other pathways.

Among the 10 hub genes, *IL1B* (31), *IL6*, *IL8* related with pro-inflammation (10, 11), *ICAM1* that regulates intercel-

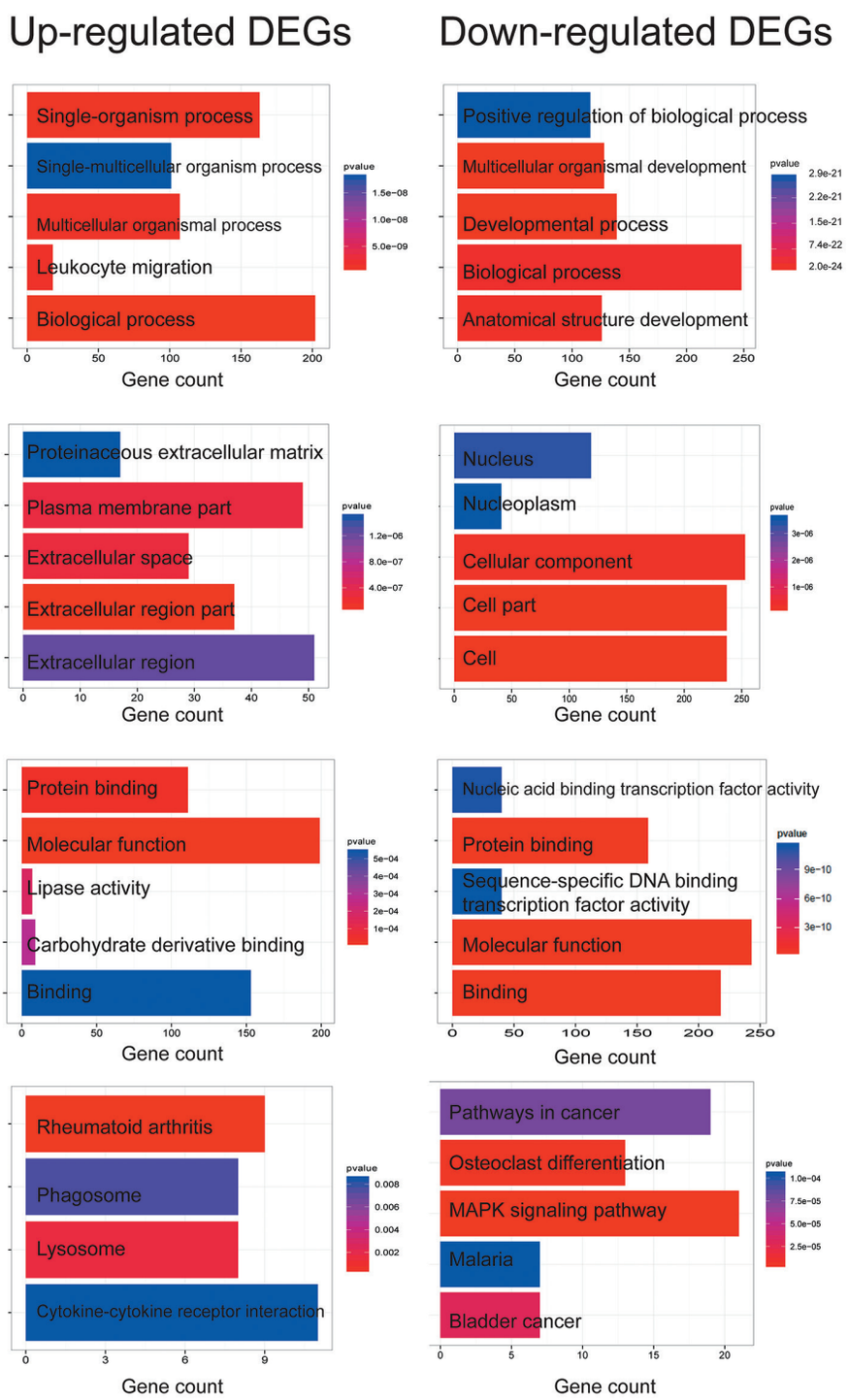

Fig. 4. The result of enrichment analysis. DEGs: differentially expressed genes.

lular adhesion (32), *ATF3* associated with neuropathic pain (33), *EGR1* that regulates cell proliferation, differentiation, and apoptosis (34), and *CDKN1A* that regulates cell cycle (35) have been reported to participate in the development of OA. While, there is little direct evidence that *VEGFA*, *JUN*, or *JUNB* was associated with OA. *VEGFA* codes vascular endothelial growth factor A. The mRNA expression of three *VEGF* isoforms (*VEGF121*, *VEGF165*, and *VEGF189*) and their receptors are elevated in OA than NC cartilage samples, and VEGF participates the destruction of OA articular cartilage through enhancing the expression of MMPs (36). As an isoform of *VEGF*, *VEGFA* might have similar bio-function. In addition, it's reported that *VEGFA* is the target gene of several microRNAs overexpressed under the OA condition (14). In this study, the expression value of *VEGFA* was significantly dysregulated, indicating that *VEGFA* might play a role in OA through regulating vascularisation of articular tissue. *JUN* and *JUNB* are proto-oncogenes and transcription factors. In this study, they were enriched in osteoclast differentiation, while *JUN* was enriched in rheumatoid arthritis as well. Reportedly, osteoclast differentiation, enhanced osteoclast activity, and cartilage degeneration are associated with subchondral osteosclerosis, which is a pathological hallmark of OA (37). In this study, *JUN* and *JUNB* were significantly dysregulated. Therefore, *JUN* and *JUNB* might be involved in the development of OA by regulating osteoclast differentiation. Among the passively changed DEGs in perturbation sub-network, *PISD* codes phosphatidylserine decarboxylase and participates in primary metabolic process, like glycine, serine and threonine metabolism and glycerophospholipid metabolism. *RARRES3*, also called tazarotene-induced gene 3 (TIG3) or retinoid-inducible gene 1 (RIG1), codes retinoic acid receptor responder protein 3. As a member of HREV107 type II tumour suppressor family, *RARRES3* induces cellular apoptosis at the Golgi apparatus in HtTA cervical cancer cells (38), and it shows decreased expression in damaged cartilage of OA patients (39). *EIF4G1* codes eukaryotic translation initiation factor 4 γ1, which regulates translation through interacting with poly(A)-binding protein (40). *EPHA3* codes ephrin receptor A3. It's reported that the *in vitro* activation of EPHB4 affects human OA articular cell metabolism, and *in vivo* overexpression of *EPHB4* has a protective effect on OA joint structural changes through reducing aggrecan, type II collagen degradation, type X collagen, and collagen fibril disorganisation (41). As a isoform of *EPHB4*, *EPHA3* might have similar bio-function. In conclusion, although there are little previous evidence that *PISD*, *RARRES3*, *EIF4G1*, and *EPHA3* participate in OA progres-

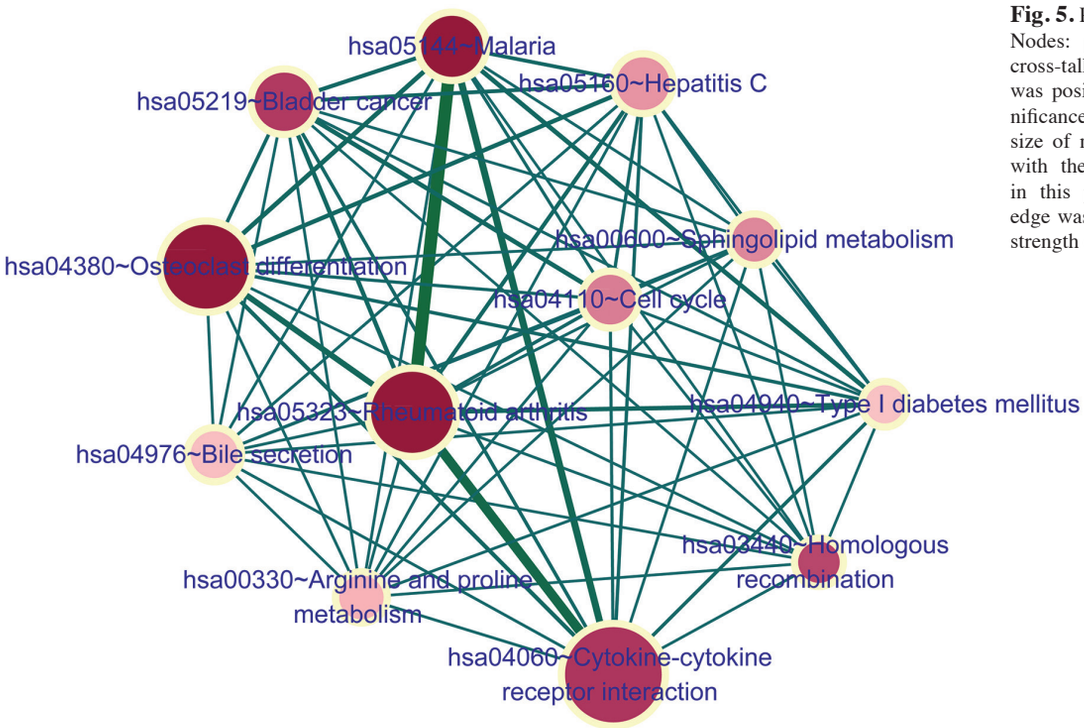

Fig. 5. Pathway cross-talk network. Nodes: pathways; edges: pathway cross-talks; the depth of node colour was positively related with the significance of pathway cross-talk; the size of node was positively related with the count of genes enriched in this pathway; the thickness of edge was positively related with the strength of cross-talk.

sion, the expression values of these genes were differentially expressed between OA and NC samples and significantly affected by the perturbation of hub genes, indicating that these genes might play a role in the development of OA.

In this study, microarray meta-analysis was performed to identify that pathways, including rheumatoid arthritis, osteoclast differentiation, and cytokine-cytokine receptor interaction, might play critical roles in the development of OA, and previously unreported genes, including *VEGFA*, *JUN*, *JUNB*, *PISD*, *RARRES3*, *EIF4G1*, and *EPHA3* might participate in OA progression, providing novel directions for drug targeting. Our further researches will focus on verifying these predictions with a plenty of clinical samples.

References

1. COOPER C, DENNISON E, EDWARDS M, LITWIC A: Epidemiology of osteoarthritis. *Medicographia* 2013; 35: 145-51.
2. VAN MANEN MD, NACE J, MONT MA: Management of primary knee osteoarthritis and indications for total knee arthroplasty for general practitioners. *J Am Osteopath Assoc* 2012; 112: 709-15.
3. IAGNOCCO A: Ultrasound in osteoarthritis. *Clin Exp Rheumatol* 2014; 32: S48-52.
4. COGGON D, READING I, CROFT P, MCLAREN M, BARRETT D, COOPER C: Knee osteoarthritis and obesity. I*nt J Obes Relat Metab Disord* 2001; 25: 622-7.
5. MA H-L, BLANCHET T, PELUSO D, HOPKINS B, MORRIS E, GLASSON S: Osteoarthritis severity is sex dependent in a surgical mouse model. *Osteoarthritis Cartilage* 2007; 15: 695-700.
6. MARTEL-PELLETIER J, LAJEUNESSE D, PELLETIER J-P: Etiopathogenesis of osteoarthritis. *Arthritis and Allied Conditions: a Textbook of Rheumatology* 2005; 15: 2199-226.
7. SPECTOR TD, MACGREGOR AJ: Risk factors for osteoarthritis: genetics. *Osteoarthritis Cartilage* 2004; 12: 39-44.
8. KERKHOF HJ, LORIES RJ, MEULENBELT I *et al*.: A genome–wide association study identifies an osteoarthritis susceptibility locus on chromosome 7q22. *Arthritis Rheum* 2010; 62: 499-510.
9. SELLAM J, BERENBAUM F: The role of synovitis in pathophysiology and clinical symptoms of osteoarthritis. *Nat Rev Rheumatol* 2010; 6: 625-35.
10. KAPOOR M, MARTEL-PELLETIER J, LAJEUNESSE D, PELLETIER J-P, FAHMI H: Role of proinflammatory cytokines in the pathophysiology of osteoarthritis. *Nat Rev Rheumatol* 2010; 7: 33-42.
11. VALCAMONICA E, CHIGHIZOLA CB, COMI D *et al*.: Levels of chemerin and interleukin 8 in the synovial fluid of patients with inflammatory arthritides and osteoarthritis. *Clin Exp Rheumatol* 2014; 32: 243-50.
12. GOLDRING MB, OTERO M: Inflammation in osteoarthritis. *Curr Opin Rheumatol* 2011; 23: 471.
13. POLUR I, LEE PL, SERVAIS JM, XU L, LI Y: Role of HTRA1, a serine protease, in the progression of articular cartilage degeneration. *Histol Histopathol* 2010; 25: 599.
14. CUADRA VMB, GONZ LEZ-HUERTA NC, ROMERO-C RDOBA S, HIDALGO-MIRANDA A, MIRANDA-DUARTE A: Altered expression of circulating MicroRNA in plasma of patients with primary osteoarthritis and in silico analysis of their pathways. *PloS one* 2014; 9: e97690.
15. BARRETT T, WILHITE SE, LEDOUX P *et al*.: NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res* 2013; 41: D991-D5.
16. UNGETHUEM U, HAEUPL T, WITT H *et al*.: Molecular signatures and new candidates to target the pathogenesis of rheumatoid arthritis. *Physiol Genomics* 2010; 42A: 267-82.
17. HUBER R, HUMMERT C, GAUSMANN U *et al*.: Identification of intra-group, inter-individual, and gene-specific variances in mRNA expression profiles in the rheumatoid arthritis synovial membrane. *Arthritis Res Ther* 2008; 10: R98.
18. HASTIE T, TIBSHIRANI R, NARASIMHAN B, CHU G: Impute: imputation for microarray data. *Bioinformatics* 2001; 17: 520-5.
19. BOLSTAD B: preprocess Core: A collection of pre-processing functions. *R package version* 2013; 1.
20. SHABALIN AA, TJELMELAND H, FAN C, PEROU CM, NOBEL AB: Merging two gene-expression studies via cross-platform normalization. *Bioinformatics* 2008; 24: 1154-60.
21. SMYTH GK: *Limma: linear models for microarray data*, in *Bioinformatics and computational biology solutions using R and Bioconductor*. 2005, Springer. p. 397-420.
22. LI J, WITTEN DM, JOHNSTONE IM, TIBSHIRANI R: Normalization, testing, and false discovery rate estimation for RNA-sequencing data. *Biostatistics* 2011: kxr031.
23. DEZA M, DEZA E: *Encyclopedia of Distances*, in *Encyclopedia of distances*. 2009, Springer Berlin Heidelberg. p. 94.

24. SZEKELY GJ, RIZZO ML: Hierarchical clustering via joint between-within distances: extending Ward's minimum variance method. *J Classif* 2005; 22: 151-83.
25. FRANCESCHINI A, SZKLARCZYK D, FRANKILD S *et al*.: STRING v9. 1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* 2013; 41: D808-D15.
26. SMOOT ME, ONO K, RUSCHEINSKI J, WANG P-L, IDEKER T: Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 2011; 27: 431-2.
27. LI F, LI P, XU W, PENG Y, BO X, WANG S: PerturbationAnalyzer: a tool for investigating the effects of concentration perturbation on protein interaction networks. *Bioinformatics* 2010; 26: 275-7.
28. YU G, WANG L-G, HAN Y, HE Q-Y: Cluster Profiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012; 16: 284-7.
29. PHAM L, CHRISTADORE L, SCHAUS S, KOLACZYK ED: Network-based prediction for sources of transcriptional dysregulation using latent pathway identification analysis. *Proc Natl Acad Sci* 2011; 108: 13347-52.
30. HARDEN JJ: A bootstrap method for conducting statistical inference with clustered data. *State Politics & Policy Quarterly* 2011; 11: 223-46.
31. SMITH A, KEEN L, BILLINGHAM M *et al*.: Extended haplotypes and linkage disequilibrium in the IL1R1–IL1A–IL1B–IL1RN gene cluster: association with knee osteoarthritis. *Genes Immun* 2004; 5: 451-60.
32. LAVIGNE P, BENDERDOUR M, LAJEUNESSE D, SHI Q, FERNANDES JC: Expression of ICAM-1 by osteoblasts in healthy individuals and in patients suffering from osteoarthritis and osteoporosis. *Bone* 2004; 35: 463-70.
33. IVANAVICIUS SP, BALL AD, HEAPY CG, WESTWOOD FR, MURRAY F, READ SJ: Structural pathology in a rodent model of osteoarthritis is associated with neuropathic pain: increased expression of ATF-3 and pharmacological characterisation. *Pain* 2007; 128: 272-82.
34. WANG F-L, CONNOR J, DODDS R *et al*.: Differential expression of egr-1 in osteoarthritic compared to normal adult human articular cartilage. *Osteoarthritis Cartilage* 2000; 8: 161-9.
35. ZHANG B, XIE Q, QUAN Y, PAN X: Expression profiling based on graph-clustering approach to determine osteoarthritis related pathway. *Eur Rev Med Pharmacol Sci* 2013; 17: 2097-102.
36. ENOMOTO H, INOKI I, KOMIYA K *et al*.: Vascular endothelial growth factor isoforms and their receptors are expressed in human osteoarthritic cartilage. *Am J Pathol* 2003; 162: 171-81.
37. GEURTS J, PATEL A, PIPPENGER BE *et al*.: Spatial association of subchondral osteosclerosis with enhanced marrow immune cell infiltration, osteoclast activity and cartilage degeneration in human osteoarthritis. *Osteoarthritis Cartilage* 2014; 22: S127-S8.
38. TSAI F-M, SHYU R-Y, JIANG S-Y: RIG1 suppresses Ras activation and induces cellular apoptosis at the Golgi apparatus. *Cell Signal* 2007; 19: 989-99.
39. SNELLING S, ROUT R, DAVIDSON R *et al*.: A gene expression study of normal and damaged cartilage in anteromedial gonarthrosis, a phenotype of osteoarthritis. *Osteoarthritis Cartilage* 2014; 22: 334-43.
40. IMATAKA H, GRADI A, SONENBERG N: A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J* 1998; 17: 7480-9.
41. VALVERDE-FRANCO G, PELLETIER JP, FAHMI H *et al*.: *In vivo* bone-specific EphB4 overexpression in mice protects both subchondral bone and cartilage during osteoarthritis. *Arthritis Rheum* 2012; 64: 3614-25.