Contents lists available at ScienceDirect

Acta Orthopaedica et Traumatologica Turcica

journal homepage: https://www.elsevier.com/locate/aott
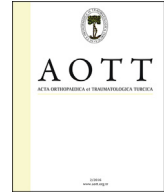
Investigation of candidate genes for osteoarthritis based on gene expression profiles

Shuanghai Dong, Tian Xia, Lei Wang, Qinghua Zhao, Jiwei Tian*

Shanghai First People's Hospital, Shanghai, China

ARTICLE INFO

ABSTRACT

Objective: To explore the mechanism of osteoarthritis (OA) and provide valid biological information for further investigation.

Methods: Gene expression profile of GSE46750 was downloaded from Gene Expression Omnibus database. The Linear Models for Microarray Data (limma) package (Bioconductor project, http://www.bioconductor.org/packages/release/bioc/html/limma.html) was used to identify differentially expressed genes (DEGs) in inflamed OA samples. Gene Ontology function enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment analysis of DEGs were performed based on Database for Annotation, Visualization and Integrated Discovery data, and protein–protein interaction (PPI) network was constructed based on the Search Tool for the Retrieval of Interacting Genes/Proteins database. Regulatory network was screened based on Encyclopedia of DNA Elements. Molecular Complex Detection was used for sub-network screening. Two sub-networks with highest node degree were integrated with transcriptional regulatory network and KEGG functional enrichment analysis was processed for 2 modules.

Results: In total, 401 up- and 196 down-regulated DEGs were obtained. Up-regulated DEGs were involved in inflammatory response, while down-regulated DEGs were involved in cell cycle. PPI network with 2392 protein interactions was constructed. Moreover, 10 genes including Interleukin 6 (IL6) and Aurora B kinase (AURKB) were found to be outstanding in PPI network. There are 214 up- and 8 down-regulated transcription factor (TF)-target pairs in the TF regulatory network. Module 1 had TFs including SPI1, PRDM1, and FOS, while module 2 contained FOSL1. The nodes in module 1 were enriched in chemokine signaling pathway, while the nodes in module 2 were mainly enriched in cell cycle.

Conclusion: The screened DEGs including IL6, AGT, and AURKB might be potential biomarkers for gene therapy for OA by being regulated by TFs such as FOS and SPI1, and participating in the cell cycle and cytokine–cytokine receptor interaction pathway.

Osteoarthritis (OA), also known as degenerative arthritis, is a type of joint disease that results from breakdown of joint cartilage and underlying bone.[1] The most common symptoms of this disease are joint pain and stiffness.[2] So far, lifestyle modifications such as weight loss and exercise, and analgesics are the mainstay of treatment, but they only provide improvement in physical function.[3] Therefore, it is important to research the mechanism of OA.

Epidemiological mechanism of OA is complex and has multiple factors including genetic, biological, and biomechanical components.[4] Recently, molecular mechanism of OA has been studied. Interleukin 1 (IL-1) has been confirmed as a critical factor that can induce joint synovitis degeneration and articular cartilage degradation.[5] In addition, tumor necrosis factor-alpha has been proven to activate multinucleated cells and stimulate synovial cells to produce prostaglandin E2 (PGE2), further attacking cartilage and bone.[6] Synovial membrane (SM) is traditionally considered part of the joint capsule.[7] The coexisting protein metabolism in SM has been confirmed to be involved in degradation of articular cartilage.[8] Decreased expression of protein in T cells from SM of OA patients is related to the progress of OA.[9]

* Corresponding author. Department of Orthopaedics, Shanghai First People's Hospital, Shanghai Jiao Tong University School of Medicine, No. 100 Haining Rd, Shanghai 200080, China.

E-mail address: dale8188@163.com (J. Tian).

Bioinformatics analysis of gene expression profiles is now providing new opportunities to uncover potential disease-related genes of SM associated with progression of OA.[10,11] By investigating the gene expression profiles of SM samples from early-stage and end-stage OA patients, genes related to immune response, cartilage development, protein glycosylation, and muscle development have been revealed to be useful in diagnosis and treatment of OA.[10] One study on gene expression patterns of synovial cells from inflamed and normal/reactive areas of SM demonstrated that genes matrix metalloproteinase 3 (MMP3) and hyaluronan synthase 1 (HAS1) were significantly up-regulated and in pathways such as cartilage catabolism pathway.[11] Although there is a close relationship between OA and SM alteration, potential disease-related genes of OA based on SM samples are still unclear.

In order to research more molecular mechanisms of OA, a bioinformatics analysis was performed in this study to explore genes or proteins that could be potential targets in OA treatment based on gene expression profile.

Patients and methods

Samples

Gene expression profile data of GSE46750[11] were downloaded from Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) based on platform GPL10558 Illumina HumanHT-12 v4 Expression BeadChip (Illumina, Inc., San Diego, CA, USA). Inflammation status of SM was classified as normal/reactive or inflamed.[11] A total of 24 samples derived from 12 knee OA patients were used for microarray data: 12 synovial tissue samples from inflamed areas of SM, and 12 from normal/reactive areas of SM in the same patients.

Data preprocessing and differential expression analysis

Normalization of this data was performed using Robust Multichip Average (RMA) method[12] of "affy − analysis of Affymetrix GeneChip data at the probe level" software package (Bioconductor project, http://bioconductor.org/packages/release/bioc/html/affy.html)[13] and R statistical software (version 3.0.0; R Project for Statistical Computing, https://www.r-project.org/). Differentially expressed genes (DEGs) in inflamed OA samples were then identified using Linear Models for Microarray Data (limma) package (Bioconductor project, http://www.bioconductor.org/packages/release/bioc/html/limma.html).[14] P-value was adjusted using Benjamini−Hochberg method.[15] Subsequently, adjusted P-value <0.05 and $\log_2$ fold change (FC) > 1 were selected as threshold value for DEG screening.

Gene Ontology annotation and pathway analysis

The Database for Annotation, Visualization and Integrated Discovery (DAVID; http://david.abcc.ncifcrf.gov/)[16] is a functional classification tool that provides a set of functional annotation tools for investigators to understand biological meaning behind large list of genes. Gene Ontology (GO, http://www.geneontology.org)[17] provides a common approach for functional studies of large-scale genomic or transcriptomic data, which mainly consists of biological process (BP), molecular function (MF), and cellular component (CC). In this study, GO BP function enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of DEGs were performed based on DAVID data.[16] P-value <0.05 and count >2 were considered threshold values.

Protein−protein interaction (PPI) network construction

Search Tool for the Retrieval of Interacting Genes/Proteins (STRING; http://string-db.org/) is a biological database of known and predicted protein−protein interactions.[18] Protein interactions were selected according to STRING database with combined score >0.4 and protein−protein interaction (PPI) network was constructed. Hub nodes with many interaction partners were identified in PPI network that were considered to be associated with development of OA.

Prediction of transcriptional regulatory relationship

Encyclopedia of DNA Elements (ENCODE) Consortium database (https://www.genome.gov/encode/) collects sequence-based studies in order to map functional factors including RNA-transcribed regions, transcription factor (TF) binding, and chromatin structure across the human genome.[19] Regulatory network was screened based on network data of TFs in ENCODE database. Then, transcriptional regulation relationship was screened for regulatory network construction by comparing it with PPI network.

Functional enrichment analysis of integrated sub-network.

Molecular Complex Detection (MCODE) clustering algorithm[20] was used for sub-network screening with default threshold (degree cutoff: 2; node score cutoff: 0.2; k-core: 2; max. depth: 100). Two sub-networks with highest PPI network node degree were integrated with transcriptional regulatory network. In addition, KEGG functional enrichment analysis was processed to annotate functions of TFs.

Results

Identification of DEGs

Based on threshold of P-value <0.05 and $\log_2$FC > 1, a total of 401 up-regulated DEGs and 196 down-regulated DEGs were identified.

Functional enrichment analysis of DEGs

To gain further insight into function of the genes, significant enrichment functions of DEGs were annotated using DAVID tool. As shown in Table 1, GO-BP functional enrichment analysis demonstrated that up-regulated DEGs were primarily involved in defense response (GO: 0006952; 82 DEGs) and inflammatory response (GO:

Table 1
Results of Gene Ontology functional enrichment analysis of differently expressed genes in osteoarthritis (Top 10 listed).

Category	Term	Description	Count	P-value
Up-regulated				
BP	GO:0006952	Defense response	82	3.15E−38
BP	GO:0006954	Inflammatory response	52	1.72E−27
BP	GO:0009611	Response to wound	62	2.52E−25
BP	GO:0006935	Chemotaxis	32	9.03E−20
BP	GO:0042330	Taxis	32	9.03E−20
Down-regulated				
BP	GO:0000279	M phase	45	2.23E−33
BP	GO:0007049	Cell cycle	61	2.21E−32
BP	GO:0000280	Nuclear division	37	1.99E−30
BP	GO:0007067	Mitosis	37	1.99E−30
BP	GO:0022403	Cell cycle phase	46	3.67E−30

BP: biological process; GO: Gene Ontology. P-value <0.05 and count >2 were threshold values for significant difference.

Table 2
Results of Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis of differentially expressed genes in osteoarthritis (Top 8 listed).

Category	Term	Description	Count	P value
Up-regulated	hsa04060	Cytokine–cytokine receptor interaction	30	1.34E−07
	hsa04062	Chemokine signaling pathway	21	2.33E−05
	hsa04640	Hematopoietic cell lineage	15	3.49E−06
	hsa05322	Systemic lupus erythematosus	15	1.89E−05
Down-regulated	hsa04110	Cell cycle	11	4.48E−07
	hsa04115	P53 signaling pathway	6	6.44E−04
	hsa04114	Oocyte meiosis	7	8.97E−04
	hsa04914	Progesterone-mediated oocyte maturation	5	1.19E−02

P-value <0.05 and count >2 were considered threshold values for significant difference.

0006954; 52 DEGs), while down-regulated DEGs were mainly involved in cell cycle (GO: 0007049; 61 DEGs) and M phase (GO: 0000279; 45 DEGs).

As demonstrated in Table 2, KEGG pathway analysis showed that up-regulated DEGs were primarily enriched in pathways like cytokine–cytokine receptor interaction (hsa04060; 30 DEGs) and chemokine signaling pathway (hsa04062; 21 DEGs), while down-regulated DEGs were mainly enriched in pathways such as cell cycle (hsa04110; 11 DEGs) and p53 signaling (hsa04115; 6 DEGs).

PPI network analysis

Total of 2392 protein interactions were revealed with combined score >0.4 and PPI network was constructed for additional analysis (Fig. 1). Top 10 hub nodes in PPI network are listed in Table 3. Results demonstrated that in this PPI network, IL6 (degree: 72), putative functional angiotensinogen (AGT; degree: 60), matrix metallopeptidase (MMP9; degree: 53) and c-Fos encoding gene (FOS; degree: 53) were identified as up-regulated genes, while aurora kinase B (AURKB; degree: 55), polo-like Kinase 1 (PLK1; degree: 54), baculoviral inhibitor of apoptosis (IAP) repeat-containing 5 (BIRC5; degree: 54), and cell division cycle 20 (CDC20) were identified as hub down-regulated genes.

Prediction of transcriptional regulatory relationship

Transcriptional regulatory network constructed is provided in Fig. 2. There are 214 up-regulated TF-target pairs and 8 down-regulated TF-target pairs. Among 214 up-regulated pairs, FOS regulated many targets, including ADAP2, ADORA3, APOC1, and

Table 3
Top 10 genes in protein–protein interaction network.

Gene	Degree	LogFC	Expression
IL6	72	1.825843	Up-regulation
AGT	60	2.100331	Up-regulation
AURKB	55	−1.86115	Down-regulation
PLK1	54	−1.07164	Down-regulation
BIRC5	54	−1.34396	Down-regulation
MMP9	53	2.262069	Up-regulation
FOS	53	1.524809	Up-regulation
CDC20	53	−1.16367	Down-regulation
CCNB2	52	−1.23723	Down-regulation
MAD2L1	50	−1.01054	Down-regulation

Gene represents the name (symbol) for each protein; degree represents the number of interactions for each gene.

Fig. 2. Transcription factor regulatory network. Red dot represents up-regulated gene, green dot represents down-regulated gene, red triangle represents up-regulated transcription factors (TFs), green triangle represents down-regulated TFs, gray edges represent protein interaction relationships. Deeper color represents stronger differential expression.

APOE, while SPI1 regulated the expression of ACP5, ACTG2, and ADAP2. Among 8 down-regulated pairs, FOSL1 regulated 8 targets, including NEK2, TUBB2, ANLN, and FOSL1.

Functional enrichment analysis of integrated sub-network

Integrated modules created after MCODE analysis and integrated analysis are shown in Fig. 3. Module 1 (MCODE score: 18.11) had 3 primary TFs: SPI1, PRDM1, and FOS, which mainly affected up-regulated DEGs. Module 2 (MCODE score: 17.619), with FOSL1, mainly affected down-regulated DEGs. KEGG pathway enrichment analysis was also processed. Nodes in module 1 were enriched in pathways such as chemokine signaling pathway (P = 2.04E−08), cytokine–cytokine receptor interaction (P = 3.08E−07), and neuroactive ligand–receptor interaction (P = 4.25E−06). Nodes in
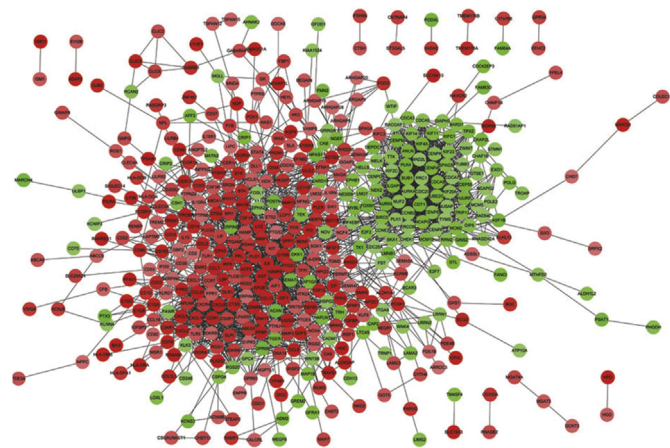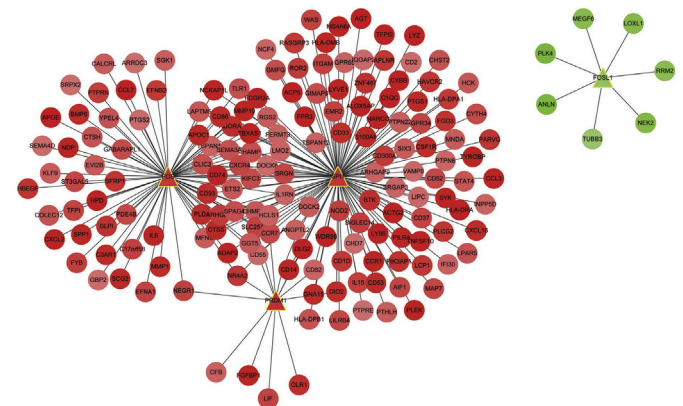

Fig. 1. Protein–protein interaction network in this study: Red represents up-regulated gene, green represents down-regulated gene, gray edges represent protein interaction relationships. Deeper color represents stronger differential expression.
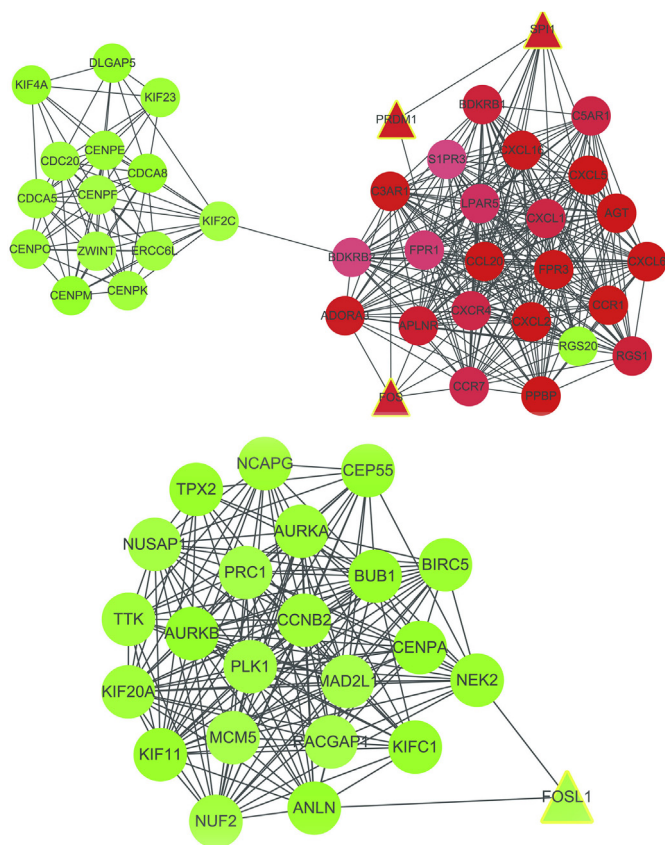
Fig. 3. Integrated modules of transcription factor regulatory network. Red dot represents up-regulated gene, green dot represents down-regulated gene, red triangle represents up-regulated transcription factors (TFs), green triangle represents down-regulated TFs, gray edges represent protein interaction relationships. Deeper color represents stronger differential expression.

module 2 were mainly enriched in pathways cell cycle (P = 4.38E−07), oocyte meiosis (P = 1.36E−05), and progesterone-mediated oocyte maturation (P = 2.46E−04).

Discussion

OA is the most common form of joint disease and the leading cause of chronic disability in middle-aged and older populations.[21] However, pathogenesis and progression of OA is still unclear. In this study, bioinformatics analysis was performed based on gene expression profile of SM samples. GO functional enrichment analysis showed that up- and down-regulated DEGs were primarily involved in defense response and cell cycle, respectively. KEGG pathway analysis showed that up- and down-regulated DEGs were mainly enriched in cytokine—cytokine receptor interaction and cell cycle, respectively. Several key genes, including IL6, AGT, and AURK, were hub nodes in PPI network.

IL6, a necessary factor for development of musculoskeletal system, has been proven to play an important role in pathogenesis of OA.[22] A previous study indicated that high production of IL6 is important immunopathological element for induction of increased B-cell response region in inflamed tissue.[23] Franckhauser et al indicated that chronically elevated IL6 levels could lead to disorder of skeletal muscles and marked inflammation in organ.[24] In this study, IL6 was revealed to be most significant up-regulated hub gene in present PPI network and regulated by FOS. Based on studies mentioned above, we speculate that IL6 may take part in the process of OA via skeletal-related disorder and inflammation.

AGT is an alpha-2-globulin that is produced constitutively and released into circulation mainly by the liver. Hong et al confirmed that angiotensin-converting enzyme (ACE) was a critical factor in physiology of vasculature, blood pressure, and inflammation, and ACE polymorphism was risk factor for severe form primary knee OA.[25] In addition, renin-angiotensin system could regulate hypertrophic differentiation of chondrocytes.[26] Furthermore, another earlier study reported that AGT precursor in synovial fluid participated in the mechanism of OA by inhibiting proteinase activities.[27] In this study, AGT was also a hub node in PPI network, regulated by SPI1. Therefore, AGT might contribute to process of OA by various pathways such as inflammation.

AURKB functions in attachment of mitotic spindle to centromere and has been identified as overexpressed in cancer.[28] A previous study proved that high expression of Aurora kinase A is associated with unfavorable cytogenetic abnormalities and increased risk of diseases.[29] In gastric cancer, AURKB is associated with low risk of cancer progression, and over-expression of AURKB can therefore be used to identify gastric cancer patients with a favorable prognosis.[30] However, relationship between AURKB and OA is rare. In the present study, AURKB was the most significant down-regulated hub gene in PPI network. This result indicates that down-regulation of AURKB may play a role in the process of OA.

Based on gene expression profiles, Zhang et al indicated that cell cycle pathway is novel therapeutic target for OA.[31] Patients with OA often exhibit inflammatory infiltrate in SM.[9] Chondrocytes, synovial cells, and other cells in the joint can express and respond to cytokines, which have been detected in synovial fluid of patients with OA.[32] From in vitro and animal model studies of OA, several cytokines have been identified as potential targets for OA therapy.[33] In present study, cytokine—cytokine receptor interaction was the most significant pathway enriched by up-regulated DEGs of OA. This result indicates that DEGs in this pathway may be vital for regulation of cytokines, which will further influence process of OA. Thus, cell cycle and cytokine—cytokine receptor interaction may be key pathways strongly associated with process of OA.

In conclusion, screened genes, such as IL6, AGT, and AURKB may be potential targets for gene therapy by being regulated by TFs such as FOS and SPI1. Moreover, key pathways such as cell cycle and cytokine—cytokine receptor interaction may be important mechanisms in OA development. Although other DEGs identified in our study have not been discussed, their roles in OA development merit further investigation. Thus, bias may exist in our study and role of these critical genes should be confirmed with experiments such as reverse transcription polymerase chain reaction and western blot.

Conflict of interest

None declared.

References

1. Glyn-Jones S, Palmer AJ, Agricola R, et al. Osteoarthritis. *Lancet*. 2015.
2. Tiernan C, Small R, Kwon Y-H, et al. Consumption of whole grape powder reduces joint pain and influences serum biomarkers in individuals with self-reported knee osteoarthritis (1025.10). *FASEB J*. 2014;28, 1025.1010.
3. Beavers D, Beavers K, Loeser R, et al. The independent and combined effects of intensive weight loss and exercise training on bone mineral density in overweight and obese older adults with osteoarthritis. *Osteoarthr Cartil*. 2014;22: 726—733.
4. Johnson VL, Hunter DJ. The epidemiology of osteoarthritis. *Best Pract Res Clin Rheumatol*. 2014;28:5—15.
5. Elsaid KA, Zhang L, Shaman Z, Patel C, Schmidt TA, Jay GD. The impact of early intra-articular administration of interleukin-1 receptor antagonist on lubricin metabolism and cartilage degeneration in an anterior cruciate ligament transection model. *Osteoarthr Cartil*. 2015;23:114—121.
6. Fahy N, Farrell E, Ritter T, Ryan AE, Murphy JM. Immune modulation to improve tissue engineering outcomes for cartilage repair in the osteoarthritic joint. *Tissue Eng Part B Rev*. 2014;21:55—66.

7. Moskalewski S, Osiecka-Iwan A, Jankowska-Steifer E, Hyc A. Synovial membrane asks for independence. *Folia Morphol*. 2014;73:395–398.
8. Fitowska A, Ostalowska A, Dobrakowski M, et al. Protein metabolism in the synovial membrane in the hip osteoarthritis. *Pol Orthop Traumatol*. 2012;77:21–26.
9. Sakkas LI, Koussidis G, Avgerinos E, Gaughan J, Platsoucas CD. Decreased expression of the CD3zeta chain in T cells infiltrating the synovial membrane of patients with osteoarthritis. *Clin Diagnost Lab Immunol*. 2004;11:195–202.
10. Ma CH, Lv Q, Cao Y, et al. Genes relevant with osteoarthritis by comparison gene expression profiles of synovial membrane of osteoarthritis patients at different stages. *Eur Rev Med Pharmacol Sci*. 2014;18:431–439.
11. Lambert C, Dubuc JE, Montell E, et al. Gene expression pattern of cells from inflamed and normal areas of osteoarthritis synovial membrane. *Arthritis Rheumatol*. 2014;66:960–968.
12. Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*. 2003;4:249–264.
13. Gautier L, Cope L, Bolstad BM, Irizarry RA. Affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*. 2004;20:307–315.
14. Diboun I, Wernisch L, Orengo CA, Koltzenburg M. Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma. *BMC Genomics*. 2006;7:252.
15. Winnepenninckx V, Lazar V, Michiels S, et al. Gene expression profiling of primary cutaneous melanoma and clinical outcome. *J Natl Cancer Inst*. 2006;98:472–482.
16. Alvord G, Roayaei J, Stephens R, Baseler MW, Lane HC, Lempicki RA. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol*. 2007;8:R183.
17. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. *Nat Genet*. 2000;25:25–29.
18. von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res*. 2003;31:258–261.
19. Raney BJ, Cline MS, Rosenbloom KR, et al. ENCODE whole-genome data in the UCSC genome browser (2011 update). *Nucleic Acids Res*. 2011;39:D871–D875.
20. Wu D, Hu X. *Topological analysis and sub-network mining of protein-protein interactions*. Idea Group Publisher; 2007.
21. Zhang M, Wang J. *Epigenet Osteoarthr Genes Dis*. 2015;2:69–75.
22. Cengic T, Trkulja V, Pavelic SK, et al. Association of TGFB1 29C/T and IL6-572G/C polymorphisms with developmental hip dysplasia: a case-control study in adults with severe osteoarthritis. *Int Orthop*. 2015;39:793–798.
23. Fujihashi K, Kono Y, Kiyono H. Effects of IL6 on B cells in mucosal immune response and inflammation. *Res Immunol*. 1992;143:744–749.
24. Franckhauser S, Elias I, Rotter Sopasakis V, et al. Overexpression of Il6 leads to hyperinsulinaemia, liver inflammation and reduced body weight in mice. *Diabetologia*. 2008;51:1306–1316.
25. Hong SJ, Yang HI, Yoo MC, et al. Angiotensin converting enzyme gene polymorphism in Korean patients with primary knee osteoarthritis. *Exp Mol Med*. 2003;35:189–195.
26. Yamagishi K, Tsukamoto I, Akagi M. Activation of the renin-angiotensin system introduces osteoarthritis. *Osteoarthr Cartil*. 2015;23:A320–A321.
27. Garner BC, Kuroki K, Stoker AM, Cook CR, Cook JL. Expression of proteins in serum, synovial fluid, synovial membrane, and articular cartilage samples obtained from dogs with stifle joint osteoarthritis secondary to cranial cruciate ligament disease and dogs without stifle joint arthritis. *Am J Vet Res*. 2013;74:386–394.
28. Bischoff JR, Anderson L, Zhu Y, et al. A homologue of *Drosophila aurora* kinase is oncogenic and amplified in human colorectal cancers. *EMBO J*. 1998;17:3052–3065.
29. Lucena-Araujo AR, de Oliveira FM, Leite-Cueva SD, dos Santos GA, Falcao RP, Rego EM. High expression of AURKA and AURKB is associated with unfavorable cytogenetic abnormalities and high white blood cell count in patients with acute myeloid leukemia. *Leukemia Res*. 2011;35:260–264.
30. Enjoji M, Iida S, Sugita H, et al. BubR1 and AURKB overexpression are associated with a favorable prognosis in gastric cancer. *Mol Med Rep*. 2009;2:589–596.
31. Zhang B, Xie QG, Quan Y, Pan XM. Expression profiling based on graph-clustering approach to determine osteoarthritis related pathway. *Eur Rev Med Pharmacol Sci*. 2013;17:2097–2102.
32. Miller RE, Miller RJ, Malfait AM. Osteoarthritis joint pain: the cytokine connection. *Cytokine*. 2014;70:185–193.
33. Blom AB, van der Kraan PM, van den Berg WB. Cytokine targeting in osteoarthritis. *Curr drug targets*. 2007;8:283–292.