European Review for Medical and Pharmacological Sciences 2013; 17: 2097-2102

Expression profiling based on graph-clustering approach to determine osteoarthritis related pathway

B. ZHANG, Q.-G. XIE, YI QUAN, X.-M. PAN

Department of Orthopedics, Chengdu Military General Hospital, Chengdu, Republic of China

Abstract. – BACKGROUND: Osteoarthritis (OA) is the most common disease of joints in adults around the world. Current available drugs to treat osteoarthritis are predominantly directed towards the symptomatic relief of pain and inflammation but they do little to reduce joint destruction. Effective prevention of the structural damage must be a key objective of new therapeutic approaches. Therefore, it is worthwhile to search for important molecular markers that hold great promise for further treatment of patients with osteoarthritis.

AIM: In this study, we used a graph-clustering approach to identify gene expression profiles that distinguish OA patients from normal samples.

MATERIALS AND METHODS: We performed a comprehensive gene level assessment of osteoarthritis using five osteoarthritis samples and five normal samples graph-clustering approach.

RESULTS: The results showed that TNFAIP3, ATF3, PPARG, etc, have related with osteoarthritis. Besides, we further mined the underlying molecular mechanism within these differently genes.

CONCLUSIONS: The results indicated tyrosine metabolism pathway and cell cycle pathway were two significant pathways, and there was evident to demonstrate them based on previous reports. We hope to provide insights into the development of novel therapeutic targets and pathways.

Key Words:
Osteoarthritis, Graph-cluster, Molecular markers.

Introduction

Osteoarthritis (OA) is a chronic degenerative joint disorder of a high prevalence that remains the leading cause of disability in aged people[1]. This disease is characterized by softening, splitting, and fragmentation of articular cartilage accompanied by subchondral bone sclerosis, bone cysts and osteophyte formation[2].

Commonly, the pathophysiology progression of OA is divided into three stages[3]. Stage I is the break-down of articular cartilage. Chondrocytes constitute the unique cellular component of articular cartilage, which synthesize the components of extracellular matrix, including collagens, proteoglycans and non-collagen proteins. The primary cause of this process is thought to be increased proteolytic enzyme activity, such as matrix metalloproteinases (MMP-1, MMP-8, and MMP-13), aggrecanases (ADAMTS-4, ADAMTS-5, and ADAMTS-9), and cathepsin[4,5]. Stage II is fibrillation and erosion of cartilage surface, accompanied by the release of breakdown products into the synovial fluid, which can promote synovial inflammation, that is, stage III. Pro-inflammatory cytokines produced by the synovium and chondrocytes, especially interleukin IL-1 and tumor necrosis factor alpha (TNF-α), play a significant role in this stage to decrease proteoglycan collagen synthesis and increasing aggrecan release. IL-1 and TNF-α also induce chondrocytes and synovial cells to produce other inflammatory mediators, such as IL-8, IL-6, nitric oxide, and prostaglandin E_2[6-8].

As an effective method, DNA microarray analysis have been extensively used to study global changes in gene expression in disease, model systems and in response to drug treatment[9]. One microarray experiment has been designed to analyze genetic expression patterns and identify potential genes to target for OA[10]. In this study, we used a graph-clustering approach to identify gene expression profiles that distinguish OA patients from normal samples. Furthermore, the relevant Geneo Ontology (GO) terms in the network was analyzed to explain potential mechanisms in response to OA.

Materials and Methods

Affymetrix Microarray Data and Differentially Expressed Genes (DEGs) Analysis

Graph-clustering approach was performed between five OA patient samples and five control donors. The GSE1919 expression profile was ob-

Corresponding Author: Bo Zhang, MD; e-mail: zhangbodr@gmail.com

tained from a public functional genomics data repository GEO (http://www.ncbi.nlm.nih.gov/geo/) which are based on the Affymetrix Human Genome U95A Array.

Statistical Analysis

For the GSE1919 dataset, the limma method[11] was used to identify DEGs. The original expression datasets from all conditions were processed into expression estimates using the robust microchip averaging (RMA) method with the default settings implemented in Bioconductor, and then construct the linear model. The DEGs only with the fold change > 2 and p-value < 0.05 were selected.

For demonstrating the potential connection, the Spearman rank correlation (r) was used for comparative target genes correlations. The significance level was set at r > 0.9 and local false discovery rate (fdr)[12] < 0.05. All statistical tests were performed with the R program (http://www.r-project.org/).

Network Analyses and Graph Clustering

To identify co-expressed groups we used DPClus (detection of protein complexes cluster)[13]. A graph clustering algorithm that can extract densely connected nodes as a cluster. It is based on density-and periphery tracking of clusters. DPClus is freely available from http://kanaya.naist.jp/DPClus/. In this study, we used the overlapping-mode with the DPClus settings. We set the parameter settings of cluster property cp; density values were set to 0.5[14] and minimun cluster size was set to eight.

Geno Ontology (GO), Interpro and Pathway Enrichment Analysis

The Gene Ontology[15] project is a major bioinformatics initiative with the aim of standardizing the representation of gene and gene product attributes across species and databases.

The InterPro, an integrated documentation resource of protein families, domains and functional sites, was created to integrate the major protein signature databases[16]. The pathway[17] database records networks of molecular interactions in the cells, and variants of them specific to particular organisms (http://www.genome.jp/kegg/).

The DAVID[18] was used to identify over-represented GO terms in biological process and pathways. The p-value < 0.01 is as the threshold for the analysis using the hypergeometric distribution.

Results

Differently Genes Selection and a Correlation Network Construction

We obtained publicly available microarray dataset GSE1919 from GEO. After microarray analysis, 178 differentially expressed genes (DEGs) with the fold change > 2 and p-value < 0.05 were selected.

To get the relationships among DEGs, r > 0.8 and fdr < 0.05 were chosen as the cutoff. Finally, 648 relationships among 161 DEGs were constructed a correlation network. The correlation network of the 161 DEGs was depicted in Figure 1. In the graph, we find that TNFAIP3, ATF3, PPARG the 3 genes have the high correlation with the r > 0.9 (not displayed in the Figure 1) between each other.

Graph Clustering Identifies Modules Significantly Enriched for DEGs Contained in Interpro Domains

At r > 0.9, DPClus[13] identified 10 clusters in the correlation network for OA; they ranged in size from 8 to 21 DEGs. Part of graph clustering results is presented in Figure 2. The clusters obtained with the graph clustering method involved the enriched domain and included the domain related with Zinc finger (p = 0.01), NADP (p = 0.07), bZIP (p = 1.08E-5) domain and Winged helix repressor (p = 0.0025) domain (Figure 4). TNFAIP3, enriched in the Zinc finger domain (not show in the Figure 2), was both in the clus1 and clus3.

GO and Pathway Enrichment Analysis of the Correlation Network in OA

To assess the significance of the clusters we used the over-represented GO terms and KEGG pathways in the clusters. Enrichment analysis by using the hypergeometrical distribution is to find the significant GO terms and pathways. Some of GO terms were enriched among these genes in the correlation network, including response to hormone stimulus, defense response, regulation of cell cycle, etc. (Table I). proto-oncogene serine/threonine protein kinase (PIM1), which enriched in the clus3 and clus8, may work in the regulation of cell cycle.

And significant pathways, such as tyrosine metabolism, cell cycle and acute myeloid leukemia were detected in the Table II. PIM1 not only works in the regulation of cell cycle but in the acute myeloid leukemia pathway.

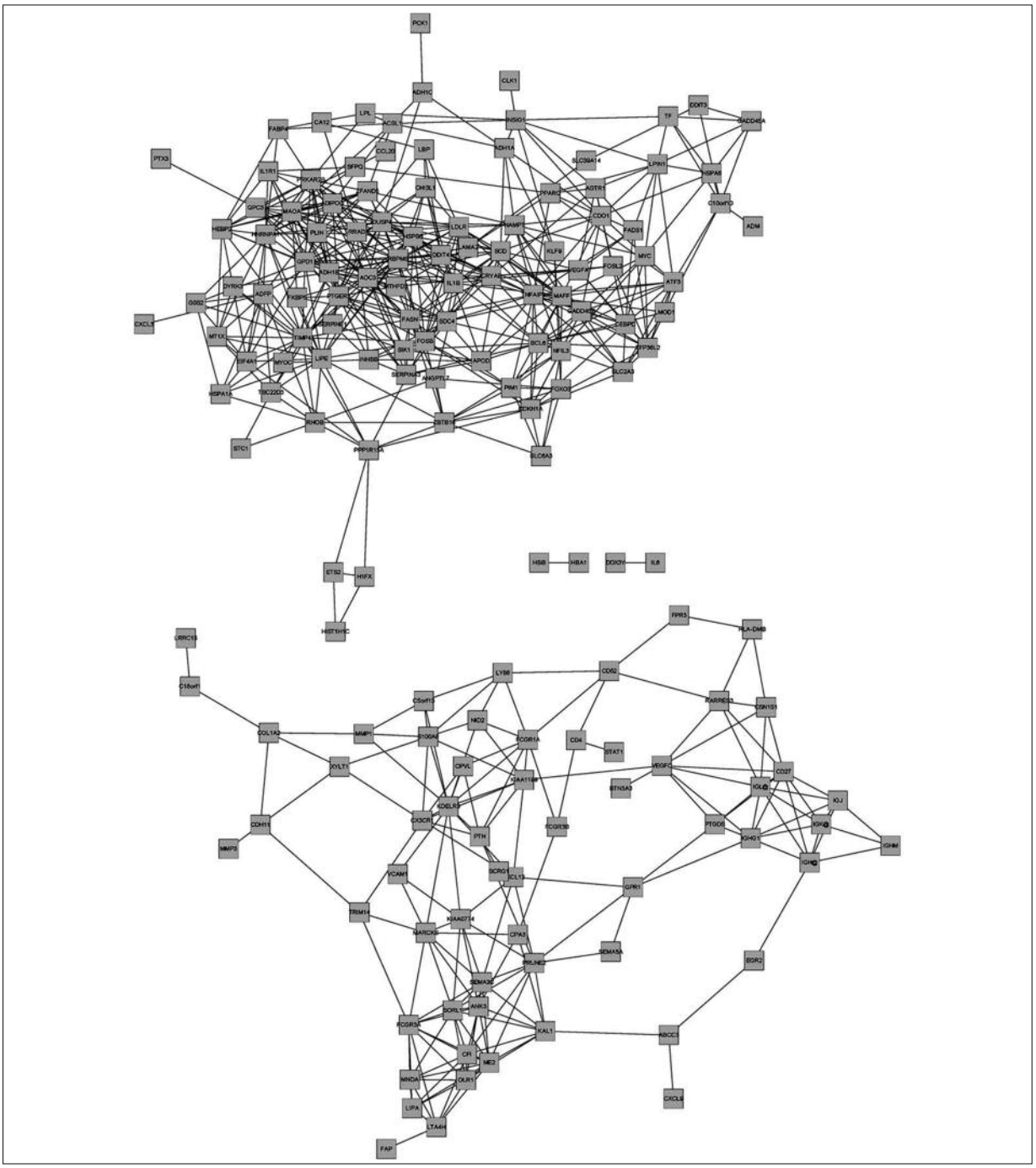

Figure 1. Correlation network of OA. Total 648 relationships with the r > 0.8 & fdr < 0.05 in the correlation network. The nodes strand for DEGs and the links strand for the high correlation among the DEGs.

Discussion

According to our analysis results, we could find that many target genes and pathways closely related with OA had been linked by our graph-clustering method (Tables I and II). Among them, TNFAIP3, ATF3, and PPARG gene have been demonstrated playing important roles in OA based on previous reports. The detail discussion was as following.

TNFAIP3, also known as zinc finger protein A20, is a dual ubiquitin-editing enzyme whose expression is rapidly induced by the tumor necrosis factor (TNF) and has been shown to involve in the negative feedback regulation of NF-κB as

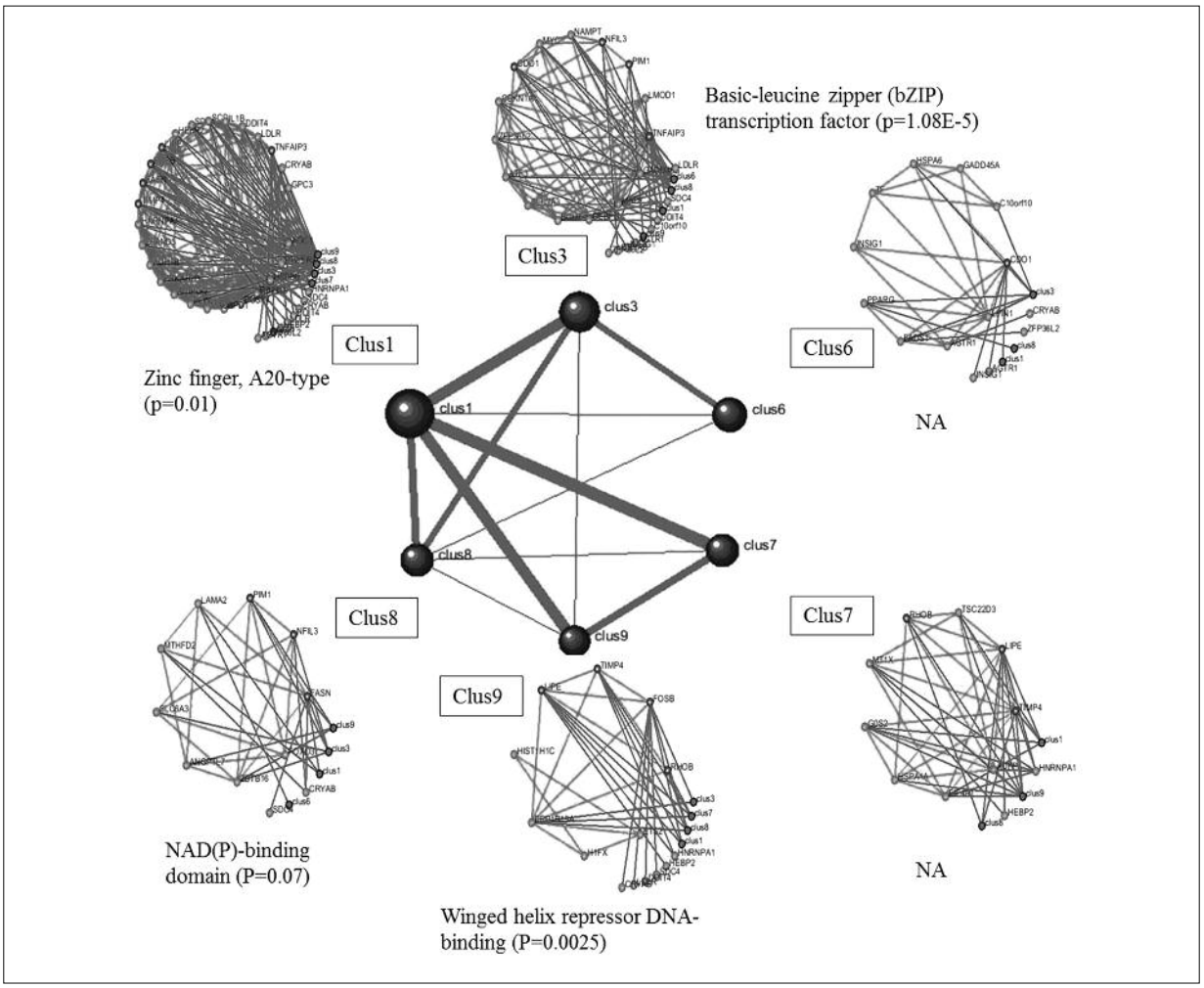

Figure 2. Graph clustering of correlated modules in OA (threshold r ≥ 0.9). Using the DPClus algorithm we extracted 10 clusters in OA (only 6 clusters were displayed). The significant domains were assigned by Interpro enrichment analysis (see Methods). The internal nodes of the clusters are connected by green edges; neighboring clusters are connected by red edges. NA means no term was detected.

well as TNF-mediated apoptosis[19]. TNFAIP3 was found heavy expression in nuclear and cytoplasmic in OA synoviocytes. Mutations in the zinc finger domains are able to disrupt the localisation of TNFAIP3 to an endocytic membrane compartment[20].

ATF3 gene encodes a member of the mammalian activation transcription factor/cAMP responsive element-binding (CREB) protein family of transcription factors. This gene is induced by a variety of signals, and is involved in the complex process of cellular stress response. OA is associated with neuropathic pain. As a marker of nerve injury, ATF-3 showed significantly increased ATF-3-immunoreactivity following MIA-induced OA treatment in lumber (L) 4 and L5 dorsal root ganglia of the ipsilateral knee[21].

PPARG (peroxisome proliferator-activated receptor gamma) is a ligand-activated transcription factor and member of the nuclear receptor superfamily. PPARG expression level is reduced in OA chondrocytes, but increase expression of inflammatory and catabolic factors, such as IL-1, TNF-α, IL-17, and prostaglandin (PG) E_2. Thus, inhibition of PPARG expression in chondrocytes by proinflammatory cytokines may be an important process in OA pathophysiology. Agonists of PPARG also have been demonstrated to inhibit inflammation and reduce synthesis of cartilage degradation products, and reduce the development/progression of cartilage lesions in OA animal models[22].

To identify the relevant pathways changed in each cluster, we used the hypergeometric distribution approach on pathway level. The results showed

Table I. List of enriched GO term in cluster 1 to 10 detected by DPClus.

Category	Term	p-value$_{min}$	Genes	Benjamini
clus1	GO:0009725~response to hormone stimulus	3.21E-04	PRKAR2B, LDLR, CRYAB, IL1B, TIMP4, ADIPOQ	0.192237
clus2	GO:0006952~defense response	0.001081	LIPA, OLR1, MNDA, LTA4H, CFI	0.14786
clus3	GO:0051726~regulation of cell cycle	2.11E-04	CDKN1A, PIM1, BCL6, ADD45B, MYC	0.093776
clus4	GO:0006952~defense response	0.006394	S100A8, LY86, FCGR1A, CX3CR1	0.607983
clus5	GO:0006955~immune response	0.001223	IGHG1, IGL@, IGJ, IGH@, CD27	0.177799
clus6	GO:0010033~response to organic substance	2.08E-05	TF, FADS1, PPARG, HSPA6, CDO1, LPIN1	0.007652
clus7	NA			
clus8	GO:0048609~reproductive process in a multicellular organism	0.031356	SLC6A3, ZBTB16, FOXO3	0.999822
clus9	GO:0006334~nucleosome assembly	0.042673	HIST1H1C, H1FX	0.994426
clus10	GO:0051222~positive regulation of protein transport	0.029356	VEGFC, CD27	0.991497

NA: no term was detected.

that cell cycle, tyrosine metabolism and acute myeloid leukemia were as the significant pathway (p-value < 0.05). And there was also evident that these pathways involved in OA progression.

Several studies have provided evidence that chondrocyte activation occurs very early in OA to increase proteoglycan content of cartilage; then the turnover of cartilage matrix is enhanced, resulting in proteoglycan depletion; and, finally, chondrocytes are lost. Therefore, chondrocyte cell cycle plays an important role in OA. Cell cycle analysis showed that the proportion of activated chondrocytes in the S phase was significantly higher in Max damage OA sample than in Min damage OA sample or normal cartilage. Amounts of cell cycle related intracellular signal involve in this process, such as p53 and Bcl-2. Bcl-2 promotes cell survival, whereas p53 can arrest cell cycle[23,24].

Tyrosine metabolism mainly includes two pathways, namely, catecholamine and melanin. Among them, catecholamine was suggested involved in OA. Catecholamine (CA) implies dopamine (DA) and its metabolic products, noradrenaline (NA) and adrenaline (A). These CAs are synthesized from the amino acid L-tyrosine (L-Tyr) in a common biosynthetic pathway that uses six enzymes. Tyrosine hydroxylase (TH) is the rate-limiting enzyme for CA biosynthesis[25]. Study showed that TH+ cells were present only in OA inflamed synovial tissue to produce catecholamine-in fibroblasts, macrophages, B cells, mast cells and granulocytes to have a strong anti-inflammatory effect. Therefore, modulation of catecholamine-producing cells could be used as a new therapeutic target in OA[26].

Conclusions

We have used network analysis as a conceptual framework to explore the pathobiology of OA based on the assumption that OA is a contextual attribute of distinct patterns of interactions be-

Table II. List of enriched pathways in cluster1 to 10 detected by DPClus.

Category	Term	p-value	Genes	Benjamini
clus1	hsa00350:Tyrosine metabolism	0.008131	MAOA, ADH1B, AOC3	0.260717
clus2	NA			
clus3	hsa04110:Cell cycle	0.008429	CDKN1A, GADD45B, MYC	0.169905
clus4-7	NA			
clus8	hsa05221:Acute myeloid leukemia	0.066546	PIM1, ZBTB16	0.729752
clus9,10	NA			

NA: no term was detected. clus 4-7: clus4, clus5, clus6 and clus7.

tween multiple genes. The salient results of our study include TNFAIP3, ATF3, PPARG, cell cycle, and tyrosine metabolism pathway which all have related with OA in direct or indirect manner. Further experiments are still indispensable to confirm our predicted target genes.

Conflict of Interest

None.

References

1) Alcaraz MJ, Megías J, García-Arnandis I, Clérigues V, Guillén MI. New molecular targets for the treatment of osteoarthritis. Biochem Pharmacol 2010; 80: 13-21.
2) Valdes AM, Spector TD. The contribution of genes to osteoarthritis. Rheum Dis Clin North Am 2008; 34: 581-603.
3) Martel-Pelletier J. Pathophysiology of osteoarthritis. Osteoarthritis Cartilage 2004; 12: 31-33.
4) Goldring MB, Goldring SR. Articular cartilage and subchondral bone in the pathogenesis of osteoarthritis. Ann N Y Acad Sci 2010; 1192: 230-237.
5) Sarzi-Puttini P, Cimmino MA, Scarpa R, Caporali R, Parazzini F, Zaninelli A, Atzeni F, Canesi B. Osteoarthritis: An overview of the disease and its treatment strategies. Semin Arthritis Rheum 2005; 35: 1-10.
6) Bertrand J, Cromme C, Umlauf D, Frank S, Pap T. Molecular mechanisms of cartilage remodelling in osteoarthritis. Int J Biochem Cell Biol 2010; 42: 1594-1601.
7) Abramson SB, Attur M. Developments in the scientific understanding of osteoarthritis. Arthritis Res Ther 2009; 11: 227.
8) Samuels J, Krasnokutsky S, Abramson S. Osteoarthritis: a tale of three tissues. Bull Hosp Jt Dis 2008; 66: 244-250.
9) Clarke PA, te Poele R, Wooster R, Workman P. Gene expression microarray analysis in cancer biology, pharmacology, and drug development: Progress and potential. Biochem Pharmacol 2001; 62: 1311-1336.
10) Ungethuem U, Haeupl T, Witt H, Koczan D, Krenn V, Huber H, von Helversen TM, Drungowski M, Seyfert C, Zacher J, Pruss A, Neidel J, Lehrach H, Thiesen HJ, Ruiz P, Blass S. Molecular signatures and new candidates to target the pathogenesis of rheumatoid arthritis. Physiol Genomics 2010; 42A: 267-282.
11) Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 2004; 3: Article3.
12) Strimmer K. Fdrtool: A versatile r package for estimating local and tail area-based false discovery rates. Bioinformatics 2008; 24: 1461-1462.
13) Altaf-Ul-Amin M, Shinbo Y, Mihara K, Kurokawa K, Kanaya S. Development and implementation of an algorithm for detection of protein complexes in large interaction networks. BMC Bioinformatics 2006; 7: 207.
14) Fukushima A, Kusano M, Redestig H, Arita M, Saito K. Metabolomic correlation-network modules in arabidopsis based on a graph-clustering approach. BMC Syst Biol 2011; 5: 1.
15) Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat Genet 2000; 25: 25-29.
16) Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bradley P, Bork P, Bucher P, Cerutti L. Interpro, progress and status in 2005. Nucleic Acids Res 2005; 33: D201-205.
17) Kanehisa M. The kegg database. Novartis Found Symp 2002; 247: 91-101; discussion 101-103, 119-128, 244-152.
18) Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using david bioinformatics resources. Nat Protoc 2009; 4: 44-57.
19) Krikos A, Laherty CD, Dixit VM. Transcriptional activation of the tumor necrosis factor alpha-inducible zinc finger protein, A20, is mediated by kappa b elements. J Biol Chem 1992; 267: 17971-17976.
20) Elsby L, Orozco G, Denton J, Worthington J, Ray D, Donn R. Functional evaluation of TNFAIP3 (A20) in rheumatoid arthritis. Clin Exp Rheumatol 2010; 28: 708-714.
21) 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: 272-282.
22) Fahmi H, Martel-Pelletier J, Pelletier JP, Kapoor M. Peroxisome proliferator-activated receptor gamma in osteoarthritis. Mod Rheumatol 2011; 21: 1-9.
23) Iannone F, De Bari C, Scioscia C, Patella V, Lapadula G. Increased Bcl-2/p53 ratio in human osteoarthritic cartilage: a possible role in regulation of chondrocyte metabolism. Ann Rheum Dis 2005; 64: 217.
24) Johnson E, Charchandi A, Babis G, Soucacos P. Apoptosis in osteoarthritis: Morphology, mechanisms, and potential means for therapeutic intervention. J Surg Orthop Adv 2008; 17: 147.
25) Flatmark T. Catecholamine biosynthesis and physiological regulation in neuroendocrine cells. Acta Physiol Scand 2000; 168: 1-17.
26) Capellino S, Cosentino M, Wolff C, Schmidt M, Grifka J, Straub RH. Catecholamine-producing cells in the synovial tissue during arthritis: modulation of sympathetic neurotransmitters as new therapeutic target. Ann Rheum Dis 2010; 69: 1853-1860.