Acta Orthopaedica et Traumatologica Turcica xxx (2016) 1–5

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

Article history:
Received 30 November 2015
Received in revised form
22 February 2016
Accepted 30 April 2016
Available online xxx

Keywords:
Differentially expressed genes
Functional enrichment analysis
Osteoarthritis
Protein–protein interaction network
Synovial membrane

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).

http://dx.doi.org/10.1016/j.aott.2016.04.002

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 Multi-chip 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.

Please cite this article in press as: Dong S, et al., Investigation of candidate genes for osteoarthritis based on gene expression profiles, Acta Orthop Traumatol Turc (2016), http://dx.doi.org/10.1016/j.aott.2016.04.002