Biochemical and Biophysical Research Communications xxx (2018) 1–5

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Uncovering transcription factor and microRNA risk regulatory pathways associated with osteoarthritis by network analysis

Zhenhua Song [a, *], Chi Zhang [b], Lingxiao He [c], Yanfang Sui [a], Xiafei Lin [a], Jingjing Pan [a]

[a] Central South University, Xiangya School of Medicine Affiliated Haikou Hospital, Haikou, Hainan, China
[b] Rehabilitation Medicine Department, The Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China
[c] XuZhou Medical University, Xuzhou, Jiangsu, China

ARTICLE INFO

Article history:
Received 20 April 2018
Accepted 24 April 2018
Available online xxx

Keywords:
Osteoarthritis
Transcription factor
microRNA
Risk regulatory pathways

ABSTRACT

Osteoarthritis (OA) is the most common form of joint disease. The development of inflammation have been considered to play a key role during the progression of OA. Regulatory pathways are known to play crucial roles in many pathogenic processes. Thus, deciphering these risk regulatory pathways is critical for elucidating the mechanisms underlying OA. We constructed an OA-specific regulatory network by integrating comprehensive curated transcription and post-transcriptional resource involving transcription factor (TF) and microRNA (miRNA). To deepen our understanding of underlying molecular mechanisms of OA, we developed an integrated systems approach to identify OA-specific risk regulatory pathways. In this study, we identified 89 significantly differentially expressed genes between normal and inflamed areas of OA patients. We found the OA-specific regulatory network was a standard scale-free network with small-world properties. It significant enriched many immune response-related functions including leukocyte differentiation, myeloid differentiation and T cell activation. Finally, 141 risk regulatory pathways were identified based on OA-specific regulatory network, which contains some known regulator of OA. The risk regulatory pathways may provide clues for the etiology of OA and be a potential resource for the discovery of novel OA-associated disease genes.

1. Introduction

As a degenerative disease of articular cartilage, osteoarthritis (OA) is a major cause of joint pain and disability in the aging population [1]. Characteristic pathological changes of OA involve synovial inflammation and osteophyte formation, angiogenesis, articular cartilage degeneration. The development of inflammation have been considered to play a key role in the initiation and progression of OA [2]. Although many studies have shown that synovial tissue inflammation may be an essential etiological factor for OA, the underlying regulatory mechanisms during osteoarthritis associated with the synovial inflammation are still poorly understood [3].

There is now a large amount of evidence showing that aberrant individual genes could not cause abnormal phenotypes, but more generally multiple dysfunctional gene-gene interactions may cooperative effects during the development and progression of complex diseases [4,5]. With the advent of high-throughput experimental technologies, various types of interaction data are being identified at a genome scale such as protein-protein and protein-DNA interactions. Integrating these interaction data into a biological network is a valuable tool to facilitate studying the pathological processes of human diseases [6]. However, the complex structure of the biological interaction network obstructs the inference of new insights and the validation of interesting risk pathways. Thus, identifying disease-related functional pathways with simple and clear structure is very important [7]. The regulatory pathway often consists of direct molecular interactions, its simple and clear structure can help us to shed light on disease pathogenic mechanisms. For example, Chen et al. demonstrated that Xbp1 promotes triple-negative breast cancer by controlling the HIF1alpha pathway [8]. Using a large dataset of colorectal cancer miRNA and gene expression profiles, Silvia et al. identified a risk pathway, in which the suppressor activity of miR-182 on the ENTPD5 gene is important for tumor metastasis [9].

In this study, the differential expressed genes (DEGs) between normal and inflamed areas of OA patients were identified. Then, we integrated multiple resources including transcription and post-

* Corresponding author.
E-mail address: taipingshengshi87@163.com (Z. Song).

https://doi.org/10.1016/j.bbrc.2018.04.189
0006-291X/

transcriptional regulations to build OA-specific regulatory network. Our results showed that the OA-specific regulatory network enriched many immune response-related functions. To take a closer step than the identification of a disease network, we identified the linear risk pathways of OA through systems-level analysis of regulatory networks. Findings of this study may lead to a better understanding of OA.

2. Materials and methods

2.1. Differential expression analysis

The gene expression profile dataset GSE46750 [10] was downloaded from the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) database. This dataset contains inflammatory and matched normal areas of a synovial membrane obtained from 12 knee OA patients. For mRNA expression data, the probe sets were mapped to Entrez Gene IDs. If multiple probes mapped to the same gene, the mean expression value of these probes was calculated to represent the gene expression level. The gene expression profile was normalized using the robust multiarray average (RMA) in R affy package [11]. Totally, 20762 genes were remained for the following analysis. We used the paired *t*-test to identify significantly differential expressed genes (DEGs) from the processed profiles of OA patients. Only genes with the p-value < 0.01 and $|\log2FC| > 1$ were considered as DEGs.

2.2. Known OA-associated genes

The GeneCards database provides manually curated genes related to diverse human diseases from multiple relevant sources [12]. We obtained the known OA-associated genes from Genecards database according to all aliases and descriptions for OA, which include ‘degenerative joint disease’, ‘osteoarthrosis’ and ‘allied disorder’, ‘degenerative polyarthritis’, ‘degenerative arthritis’, ‘hypertrophic arthritis’ and ‘arthropathy’.

2.3. Construction of TF and miRNA regulatory network

The TF and miRNA regulatory network was constructed by integrating molecular interaction relationships of five databases, which include TRANSFAC [13], TransmiR [14], miRTarBase [15], miRecords [16] and TarBase [17]. All of the individual interactions between the curated human regulations (TFs, miRNAs and target genes) in the regulatory network are literature-supported.

2.4. Extraction of OA-specific regulatory network

The differential expression analysis could be a traditional method to identify the key genes associated with disease. While some studies also point out that the disease-related genes do not necessarily aberrant expressed in the disease status [18]. To further identified genes and regulations associated with OA, we hypothesized that the DEGs and their immediate neighbors in the curated TF-miRNA regulatory network potentially contributed to the pathology of OA. The DEGs mapped into the regulatory network as risk seed nodes. Then, we extracted the OA-specific TF-miRNA regulatory network using the DEGs and their immediate neighbors.

2.5. Identification of OA-specific risk regulatory pathways

For the regulatory network, we expected that a node (gene/miRNA) with a 0-indegree should located at the beginning of the regulatory pathway because it could not be regulated by other regulators. Similarly, a node (gene/miRNA) with a 0-outdegree should located at the end of the regulatory pathway because it could not regulate other genes/miRNAs. Therefore, we defined all of the linear pathways as those from beginnings whose indegree is 0 to ends whose outdegree is 0. From the OA-specific regulatory network, we extracted all linear pathways as the risk regulatory pathways.

2.6. Functional enrichment analysis

To explore the biological significance of genes contained in the OA-specific regulatory network, we used the clusterProfiler [19] bioinformatics tool to analyse Gene Ontology (GO) category enrichment. The adjusted p-values were obtained from multiple testing of Benjamini and Hochberg (BH) method. The term with adjusted p-value < 0.05 were selected as the significant function.

3. Results

3.1. DEGs in OA

Using the gene expression profiles of inflammatory and matched normal tissues of OA patients, we performed differential expression analysis through a paired *t*-test method. As a result, we detected 89 significantly differential expressed genes with 10 down-regulated and 79 up-regulated genes (Fig. 1). Among these DEGs, we found five genes (IL6, NR4A2, CXCL6, STAT4, PTGES) were known OA-associated genes curated in GeneCards.

3.2. The OA-specific regulatory network

To systematically understand the underlying molecular mechanism of OA, we constructed a comprehensive TF-miRNA regulatory network by integrating five manully curated databases, including TRANSFAC, TransmiR, miRTarBase, miRecords and TarBase, which address transcriptional and post-transcriptional regulations. This regulatory network contains 411 TFs, 387 miRNAs, 2300 target genes and 6036 regulations. We defined the 89 DEGs as the risk seed nodes and mapped them into the regulatory network to extract OA-specific regulatory network (Fig. 2). Specifically, we constructed the candidate risk regulatory network by connecting all of the risk seeds with their immediate neighbors. We found the OA-specific regulatory network was comprised of 97 interactions involving 38 TFs, 31 miRNAs and 35 protein-coding genes, in which 22 genes were differential expressed.

3.3. Characterization of OA-specific regulatory network

Examination of the in-degree and out-degree distribution of the OA-specific regulatory network revealed a power-law distribution, showing that the OA-specific regulatory network was scale-free, similar to most biological networks (Fig. 3). In addition, we performed function enrichment analysis for the OA-specific regulatory network. The significantly enriched GO terms were displayed in Fig. 3. Among the top of significantly enriched biological processes, we found many immune response related functions, such as leukocyte differentiation, myeloid differentiation and T cell activation. A number of studies have reported that abnormal innate immune inflammatory responses considerably contribute to the pathogenesis of OA [3,20]. These results also support the OA-specific regulatory network may be a potential resource for OA-related studies.

3.4. The OA-specific risk regulatory pathways

Identification of biologically important pathways from

Please cite this article in press as: Z. Song, et al., Uncovering transcription factor and microRNA risk regulatory pathways associated with osteoarthritis by network analysis, Biochemical and Biophysical Research Communications (2018), https://doi.org/10.1016/j.bbrc.2018.04.189