Analysis of intervertebral disc-related genes


Department of Orthopedic,
Changhai Hospital of the Second Military Medical University,
Shanghai, China

Corresponding author: M. Li
E-mail: limingllmm@hotmail.com

Received January 7, 2012
Accepted June 6, 2013
Published March 24, 2014
DOI http://dx.doi.org/10.4238/2014.March.24.7

ABSTRACT. Intervertebral disk disease is a common clinical disorder manifested by pain, ataxia, paresis, motor paralysis, and sensorimotor paralysis. The clinical features, diagnosis, and treatment of cervical and thoracolumbar disk disease have been unclear until now. In this study, some differentially expressed genes were identified, and a network was constructed based on these genes. Through the statistical analysis of nodes and the contrast of 2 more connectivity nodes, it was found that the nodes in the network are in an important position and play key roles. Several of these genes, including MAP2K6, MAP2K3, and MAPK14, belong to the MAP kinase family, and several genes, including RHOBTB2, RHOQ, and RHOH, belong to the RHO family. Therefore, we hypothesize that the development of intervertebral disk disease is related to MAP and RHO family proteins.

Key words: Intervertebral disk disease; MAP kinase family; Rho family
INTRODUCTION

Intervertebral disc disorder is a condition that involves deterioration, herniation, or other dysfunctions of the intervertebral disc (Danielsen et al., 2011). The human spinal column is composed of 33 vertebrae that are in stacked bone segments: 7 cervical, 12 thoracic, 5 lumbar, 5 sacral (fused), and 2 to 4 coccygeal (fused) (Gambardella, 2008). The intervertebral discs are structures that lie between the individual vertebrae and act as a cushion between the discs. When the nucleus pulposus starts to dehydrate, its ability to transfer and distribute loads between the vertebrae is being restricted, which will cause chronic pain (Savolainen et al., 1998).

The clinical features are unclear and the diagnosis and treatment of cervical and thoracolumbar disk disease are challenging. Clinical signs are determined by neuroanatomic localization and the severity of the spinal cord injury (Taveras and Wood, 1976; De Lahunta and Glass, 2008). Myelography is most commonly used in clinical diagnosis; however, its accuracy is not determined (Bolender et al., 1985). Thus, recent studies have provided a new understanding of the diagnosis and treatment of intervertebral disc disorder.

Based on intervertebral disc-related mRNA expression profile data, we obtained differentially expressed mRNAs between the third and fourth stages of the disease compared with the second stage of the intervertebral disc disease. In addition, we built two interaction networks at the center of the target genes. Finally, we used GO analysis to perform statistical analysis of the network nodes and to find the genes which may play important roles at the different stages of intervertebral disc disease.

MATERIAL AND METHODS

Expression profiles of intervertebral disc disease-related genes

To recognize differentially expressed genes of intervertebral disc disease in 3 different stages from the molecular level and the biological network level, we downloaded the GSE15277 data set from the Gene Expression Omnibus (GEO: http://www.ncbi.nlm.nih.gov/geo/) database. The experiment tested the expression of intervertebral disc in different stages, including 15 samples (5 second-stage samples, 7 third-stage samples, and 3 fourth-stage samples).

Differential expression analysis

First, we converted the original data into recognizable files and then filled in the lacking data. Next, all data were normalized using log (Ration) with the affy package of the R language (Ihaka and Gentleman, 1996). The mRNA expression profiles were used to perform differential expressional analysis, using the t-test to recognize differentially expressed genes between the third and fourth stages of the disease compared with the second stage of the disease. To verify that the P values were closely related to the fold-change, we used the volcano plot to reflect the P values directly. P values were corrected by the BH (Beyer-Hardwick) method, and a threshold P value of 0.05 was defined (Smyth, 2004).

Interaction network mapping

We numbered the downloaded probes with the ID converter and used the STRING
D.J. Wu et al.

(Search Tool for the Retrieval of Interacting Genes/Proteins) database to predict the interaction of objects of a protein encoded by a target gene. The results of the 2 network node analyses suggested the role that the target genes may play in the network built from sequence characteristics, species, and structure.

**GO analysis**

We annotated the nodes of the interactional network map using GO database (Ashburner et al., 2000; Wang et al., 2012). Finally, we identified the genes that may play key roles at the different stages of intervertebral disc disease.

**RESULTS**

**Differential expression analysis**

Due to various reasons, such as background and probe design challenges, there is a great difference in quality of the original data. Thus, we had to perform normalization of the data.

The $t$-test and BH multiple test correction were used to conduct differential expression analysis of the genes, and the threshold was set at 0.5. We selected $P < 0.05$ as the subsequent significance level for identifying differentially expressed genes (Table 1 and Figure 1).

**Table 1. Differentially expressed genes.**

<table>
<thead>
<tr>
<th>ID_REF</th>
<th>$t$-test</th>
<th>adj</th>
<th>logFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>g10835026_3p_a_at</td>
<td>1.52E-06</td>
<td>0.030373</td>
<td>0.898943</td>
</tr>
<tr>
<td>g6642743_3p_at</td>
<td>1.98E-06</td>
<td>0.030373</td>
<td>0.897715</td>
</tr>
<tr>
<td>Hs.21236.0.A1_3p_at</td>
<td>3.58E-06</td>
<td>0.043933</td>
<td>-0.48729</td>
</tr>
<tr>
<td>Hs.296702.0.S1_3p_at</td>
<td>9.03E-07</td>
<td>0.027704</td>
<td>0.821719</td>
</tr>
<tr>
<td>Hs.55879.2.A1_3p_s_at</td>
<td>4.32E-07</td>
<td>0.026507</td>
<td>-0.51696</td>
</tr>
</tbody>
</table>

In order to verify that the P value obtained by different ways and value of the fold-change (FC) were logical, we used the volcano plot (Figure 1). The smaller the P value, the greater the fold-change.

**Conversion of the probe number to the corresponding gene name**

The raw data only provided the gene number; therefore, the probe numbers were converted to the corresponding gene names (Table 2).
Table 2. Conversion of the probe number and corresponding gene name.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Ensembl gene</th>
<th>UniGene</th>
<th>EntrezGene</th>
<th>Affymetrix</th>
<th>GenBank accession</th>
<th>Ensembl* Chr</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAP2K6</td>
<td>ENSG00000108984</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ENSG00000177320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCC10</td>
<td>ENSG00000124574</td>
<td>Hs.55879</td>
<td>52</td>
<td>[U133_X3P] g10835026_3p_a_at</td>
<td>NM_033450</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Hs.55879.2.A1_3p_s_at</td>
<td>Hs.55879.2.A1_3p_s_at</td>
<td></td>
<td>BC064527</td>
<td>BQ051912</td>
<td>6</td>
</tr>
<tr>
<td>CLC1</td>
<td>ENSG000000213719</td>
<td>Hs.414565</td>
<td>1192</td>
<td>[U133_X3P] g5426566_3p_at</td>
<td>BS558439</td>
<td>11</td>
</tr>
<tr>
<td>SPAG5</td>
<td>ENSG00000076382</td>
<td>Hs.514033</td>
<td>10615</td>
<td>[U133_X3P] g5453631_3p_at</td>
<td>AF343547</td>
<td>17</td>
</tr>
<tr>
<td>DDR2</td>
<td>ENSG00000162733</td>
<td>Hs.593833</td>
<td>2731</td>
<td>[U133_X3P] g5453831_3p_at</td>
<td>BC052998</td>
<td>1</td>
</tr>
<tr>
<td>CCDC16</td>
<td>ENSG000000198783</td>
<td>Hs.413678 Hs.691125</td>
<td>91603</td>
<td>[U133_X3P] g7656841_3p_at</td>
<td>AF052957</td>
<td>17</td>
</tr>
<tr>
<td>ABR</td>
<td>ENSG00000159842</td>
<td>Hs.159306</td>
<td>29</td>
<td>[U133_X3P] Hs.10487.0.S1_3p_at</td>
<td>AK124547</td>
<td>17</td>
</tr>
<tr>
<td>MAPKAP1</td>
<td>ENSG00000119487</td>
<td>Hs.493138 Hs.597362</td>
<td>79109</td>
<td>[U133_X3P] Hs.98267.0.A1_3p_at</td>
<td>NM_0510092</td>
<td>17</td>
</tr>
</tbody>
</table>

Construction and analysis of network

First, we summarized all of the nodes of each network and selected the nodes that
connectivity warranted further analysis (Figure 2). Through statistical analysis of the nodes, and by comparing to 2 additional connectivity nodes, we found unique network nodes and high-value nodes, which indicated that these nodes were located in an important position and played key roles in the network (Table 3).

![Figure 2](image)

**Figure 2.** Network of differentially expressed genes.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Stage 2 vs Stage 3</th>
<th>Stage 3 vs Stage 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAP2K6</td>
<td>96</td>
<td>RAC2</td>
</tr>
<tr>
<td>MAP2K3</td>
<td>63</td>
<td>RHOB</td>
</tr>
<tr>
<td>MAPK14</td>
<td>63</td>
<td>RHOBTB2</td>
</tr>
<tr>
<td>JUN</td>
<td>58</td>
<td>RHOQ</td>
</tr>
<tr>
<td>TNF</td>
<td>46</td>
<td>RHOG</td>
</tr>
<tr>
<td>NFKB1</td>
<td>44</td>
<td>RHOBTB1</td>
</tr>
<tr>
<td>FAM48A</td>
<td>44</td>
<td>RHOD</td>
</tr>
<tr>
<td>MAPK11</td>
<td>41</td>
<td>RHOU</td>
</tr>
<tr>
<td>CDC42</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td></td>
<td>37</td>
<td></td>
</tr>
</tbody>
</table>

In these genes, MAP2K6, MAP2K3, MAPK14, etc. belong to the MAP kinase family, as well as RHOBTB2, RHOQ, RHOH, etc. belong to the RHO family. So we considered that the development of intervertebral disk disease is related to MAP and RHO families.

**DISCUSSION**

Intervertebral disk disease is a common clinical disorder manifested by pain, ataxia, paresis, motor paralysis, and sensorimotor paralysis (Tokuhashi et al., 2001). The literature by Hoerlein, Olsson, Hansen, Funquist, and many others, formed the foundation of our current medical and surgical therapies for intervertebral disk disease protrusion (Shores, 1981). Numerous manuscripts are related to surgical technique, radiologic diagnosis, pathology, pathophysiology, and incidence of intervertebral disk disease (Bagley, 1999). Furthermore, the proteoglycan and interstitial fluid content is known to decrease with age in the anulus fibrosus, decreasing the ability to absorb shock.

In our study, we obtained some differentially expressed genes between 2 vs 3 and 2 vs 4, including g10835026_3p_a_at, g6642743_3p_at, g4426566_3p_at, and g5453631_3p_at. Through statistical analysis of the nodes in the networks as well as comparison to 2 connectivity nodes, we found nodes with important positions that may play key roles.
Indications for surgical management are pain or paresis unresponsive to medical therapy and recurrent or progressive signs of thoracolumbar intervertebral disk disease (Gambardella, 2008). The major postoperative considerations are based on the surgeon’s experience, ability, and preference (McPhee et al., 2010). Most of these concerns are valid also for medically managed intervertebral disk disease patients.

Biochemically, intervertebral disk disease consists of proteoglycans, glycoproteins, and both collagenous and noncollagenous proteins (Diwan et al., 2000). The therapeutic measures of choice for intervertebral disk disease have long been debated (Sussman, 1975). A plethora of genes has been identified that are related to intervertebral disk disease, including genes related to intervertebral disk stability, inflammation, and signaling pathways. In our analysis, we used expression profile data to analyze differentially expressed genes at different stages of intervertebral disk disease and found that many of the key genes (nodes on the network) are relevant to signaling pathways.

Medical therapy and surgical therapy seek to alleviate the pain or neurologic deficits associated with intervertebral disk disease. Degeneration of intervertebral disk disease occurs with aging and results in significant changes in biochemical characteristics. The progress of intervertebral disc disease is regulated by a large combination of genes.

The progress of intervertebral disk disease is regulated by multiple genes. Studied and unknown genes related to intervertebral disk disease are likely involved in the occurrence and development of this disease. Therefore, it is necessary to develop long-term, large sample case-control studies of different groups to aid the study of intervertebral disc disease genetic mechanisms. These studies would allow personalized prevention and treatment of intervertebral disk disease.

RHOBTB2, RHOQ, and RHOH are members of the mitogen-activated protein (MAP) kinase family, also known as the extracellular signal-regulated kinase family. The MAP family is involved in many signaling cascades. MAP family members are phosphorylated when they respond to inflammatory cytokines or environment stress. The p38 protein, an important part of the MAP kinase-mediated pathway, participates in many cellular processes such as cell cycle arrest, transcriptional activation, and apoptosis (Johnson and Lapadat, 2002; Olek et al., 2003; Diao et al., 2010; Danielsen et al., 2011). The pathologic changes associated with acute spinal cord compression are related to dynamic force of compression.

In signaling pathways, the RHO family of proteins plays the role of a molecular switch by influencing the cytoskeleton. The RHOBTB2 proteins can participate in signal transduction pathways and regulate cell growth by affecting cell apoptosis and the cell cycle.

REFERENCES


