Analysis of key genes and pathways involved in acute lung injury in a mouse model

Q.H. Han¹*, N. Han¹*, Y.Z. Liu¹, Q.H. Jin², Q.Y. Lu¹ and Z.C. Li¹

¹Department of Emergency and Trauma Surgery, East Hospital Affiliated to Tongji University, Shanghai, China
²Internal Medicine, Longhua Community Health Center, Shanghai, China

*These authors contributed equally to this study.

Corresponding authors: Q.Y. Lu / Z.C. Li
E-mail: lunginjuryluli@hotmail.com

Received June 6, 2013
Accepted October 22, 2013
Published June 18, 2014
DOI http://dx.doi.org/10.4238/2014.June.18.1

ABSTRACT. A mouse model of acute lung injury (ALI) was chosen in this study to explore the key genes and pathways involved in the process of ALI with microarray technology. Gene expression microarray data were downloaded from the Gene Expression Omnibus database. Mice from the experimental group were further divided into 6 subgroups, which received octadecenoate treatments for 1, 1.5, 3, 4, 18, and 24 h. Differentially co-expressed genes were screened to uncover the pathogenesis of ALI. Almost all of the differentially co-expressed genes were identified at two times: 1.5 and 3 h. Functional analysis revealed that several inflammation-related pathways were significantly enriched. Ubiquitin-mediated proteolysis, hematopoietic cell lineage, and leukocyte transendothelial migration were enriched at 1.5 h. The B cell receptor signaling pathway, T cell receptor signaling pathway, natural killer cell-mediated cytotoxicity, Fc epsilon RI signaling pathway, and ubiquitin-mediated proteolysis were significantly enriched at 3 h. It could be inferred that ALI initiated at 1.5 h and lasted through 3 h. However, co-expression patterns were not found from 4 h onward. In conclusion, several key genes and pathways implicated...
in the development of ALI were found in this study using the mouse model, among which ubiquitin-mediated proteolysis appears to play an important role in the process.

**Key words:** Acute lung injury; KEGG pathway; Mouse model; Inflammation

### INTRODUCTION

Acute lung injury (ALI) is a secondary lung injury caused by inflammation involving a variety of mediators and effector cells. It is characterized by acute hypoxic respiratory insufficiency and exudative lesions in lung imaging (Bernard et al., 1994). ALI may develop into acute respiratory distress syndrome (ARDS), which carries a high mortality rate (40-50%) (Matthay et al., 2003). The longer the risk factors last, the higher the incidence of ARDS (Iribarren et al., 2000). Studies have indicated a close relationship between the inflammatory response and ALI (Dellinger et al., 2004; Calvano et al., 2005). The inflammatory response is an important part of the body’s innate immune system. There are both pro- and anti-inflammatory cytokines in the body; pro-inflammatory cytokines are necessary for initiation of the response, whereas anti-inflammatory cytokines control and “calm down” the response. The timing and balance between the two types of cytokines are important for normal function as excessive inflammatory reactions can cause self tissue damage (Gogos et al., 2000).

High-throughput microarray technology is a suitable choice for studying the mechanism of ALI. Wurfel (2007) proposed a new hypothesis of disease pathogenesis in ARDS with microarray technology. Wang et al. (2008) identified pre-elafin as a biomarker in ARDS through a genome-wide expression analysis. However, some shortcomings of this approach exist. For example, Gogos et al. (2000) used microarray data acquired from animal models suffering from both infection and inflammation, which complicates the expression profiling data and influences the analysis results.

In this study, we adopted mouse models treated with octadecenoate (OA) in order to exclude the potential impact of infection on the expression profile. Although the progression of OA-induced ALI is much faster than that of human ALI, they share similar symptoms. We analyzed microarray data collected at a series of time points to track the change in expression levels over time. Considering the fact that genes and proteins interact to exert biological functions (Han et al., 2004; Huang et al., 2007), we not only screened for differentially co-expressed genes (DCGs), but also performed a pathway analysis to better portray the process of ALI.

### MATERIAL AND METHODS

**Expression microarray data of the mouse model**

Mice treated with OA were selected as the animal model for ALI in this study. The gene expression microarray data set GSE18712 (Lesur et al., 2010), consisting of 56 samples, was downloaded from the Gene Expression Omnibus (GEO) database; 51 from the experimental group and 5 from the control group. Raw data were collected from the MusV29K platform and were quantified using the Bzscan2 software. NylonArray was applied for the
diagnosis and normalization of the data. The median was determined as the expression level for probes corresponding to the same symbol.

**Modeling process**

Twenty milliliters of OA was intravenously administrated to female mice in the experimental group, and an equivalent volume of physiological serum was administered to mice in the control group. Samples were collected at a series of times: 1, 1.5, 3, 4, 18, and 24 h. The number of samples in each subgroup is listed in Table 1, and the comparison model is illustrated in Figure 1.

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>1 h</th>
<th>1.5 h</th>
<th>3 h</th>
<th>4 h</th>
<th>18 h</th>
<th>24 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samples</td>
<td>5</td>
<td>8</td>
<td>11</td>
<td>13</td>
<td>8</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

*Table 1. Number of samples for experimental groups and the control group.*

**Screening of DCGs**

The DCGL package (Liu et al., 2010; Yu et al., 2011) of R was used to analyze the microarray data and to screen the DCGs. The DCe method was adopted and the whole screening process consisted of two steps: first, differential co-expression relationships (DCL) were determined among genes, and then genes that were sufficiently associated were identified as DCGs based on the binomial model (Yu et al., 2011). The ‘qth’ method was assigned to determine the relationship, and the cutoff value was set to 0.25. The ‘nbins’ was set to 20 and the P value was set to 0.05 in the LFC model (Mutch et al., 2002).

**Pathway-enrichment analysis**

To understand the biological functions of the DCGs, EASE (Hosack et al., 2003) (developed by the DAVID Bioinformatics team) was utilized for KEGG pathway-enrichment analysis. The annotation file was downloaded from DAVID. An EASE score <0.1 was regarded as evidence for significant enrichment.

**RESULTS**

**DCGs**

The subgroups of the experimental group were respectively compared with the control...
group to screen for DCGs. With a false-discovery rate ≤0.05 set as the cutoff value, DCGs were only identified for 1.5 and 3 h, and the values were relatively low. Extending the cutoff to P < 0.05 and P < 0.1, most of DCGs were still identified at 1.5 and 3 h (Table 2), and no or few DCGs were found at 4 h or later. DCGs screened out on the basis of P < 0.1 were further analyzed with respect to their functions.

### Table 2. Number of differentially co-expressed genes identified based upon various cutoff values.

<table>
<thead>
<tr>
<th></th>
<th>1 h</th>
<th>1.5 h</th>
<th>3 h</th>
<th>4 h</th>
<th>18 h</th>
<th>24 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDR &lt; 0.05</td>
<td>0</td>
<td>36</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P &lt; 0.05</td>
<td>0</td>
<td>178</td>
<td>81</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>P &lt; 0.1</td>
<td>0</td>
<td>256</td>
<td>206</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

FDR = false-discovery rate.

### Pathway-enrichment analysis

KEGG pathway analysis was performed for the 256 DCGs identified at 1.5 h. A total of 6 functional terms were significantly enriched at P < 0.1 (Table 3). The most significantly enriched pathway was ubiquitin-mediated proteolysis.

### Table 3. Functional analysis results for differentially co-expressed genes identified at 1.5 h.

<table>
<thead>
<tr>
<th>Term</th>
<th>Genes</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>mmu04120:Ubiquitin mediated proteolysis</td>
<td>Rchy1, Mdm2, Cdc20, Pias2, Socs3, Cal7, Cul3, Skp1a, Rbx1</td>
<td>0.003085</td>
</tr>
<tr>
<td>mmu04110:Cell cycle</td>
<td>Stag2, Pdkc, Mdm2, Cdc20, Tgb2, Skp1a, Sfn, Rbx1</td>
<td>0.00826</td>
</tr>
<tr>
<td>mmu04640:Hematopoietic cell lineage</td>
<td>Cdf5, Ilra, Mmeh2, E2fb1, Gp9, Csf1</td>
<td>0.017703</td>
</tr>
<tr>
<td>mmu04670:Leukocyte transendothelial migration</td>
<td>Igb2, Chdh6, Sip1, Vcam1, Ezr, Vcl</td>
<td>0.064141</td>
</tr>
<tr>
<td>mmu05200:Pathways in cancer</td>
<td>Traf2, Msh6, Wnt5a, Mdm2, Pias2, Jak1, Tgb2, Lamb1-1, Lef1, Rbx1, Fox</td>
<td>0.066667</td>
</tr>
<tr>
<td>mmu03040:Spliceosome</td>
<td>Cherp, Prp6f, Sfas2, Sfas5a, Smnap, Sfns3</td>
<td>0.07377</td>
</tr>
</tbody>
</table>

A total of 7 functional terms were significantly enriched for the 206 DCGs identified at 3 h (Table 4). Several terms were found to be associated with inflammation, such as the B cell receptor signaling pathway, T cell receptor signaling pathway, and natural killer cell-mediated cytotoxicity. Ubiquitin-mediated proteolysis remained significant at this time, which may suggest its importance in the inflammatory cascade.

### Table 4. Functional analysis results for differentially co-expressed genes identified at 3 h.

<table>
<thead>
<tr>
<th>Term</th>
<th>Genes</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>mmu04862:B cell receptor signaling pathway</td>
<td>Nfkbib, Cd19, Btk, Rac2, Rasgrp3, Nrass</td>
<td>0.00327</td>
</tr>
<tr>
<td>mmu04720:Long-term potentiation</td>
<td>Itpr1, Calm1, Rps6ka2, Pppl1cc, Nras</td>
<td>0.01175</td>
</tr>
<tr>
<td>mmu04660:T cell receptor signaling pathway</td>
<td>Nfkbib, Il10, Lat, Cdg247, Nras, Nck1</td>
<td>0.016704</td>
</tr>
<tr>
<td>mmu04680:Long-term potentization</td>
<td>Itgb2, Hnrr1, Lat, Cdg247, Nras</td>
<td>0.019035</td>
</tr>
<tr>
<td>mmu04644:Fc epsilon RI signaling pathway</td>
<td>Btk, Rac2, Lat, Nras</td>
<td>0.085901</td>
</tr>
<tr>
<td>mmu04620:Ubiquitin-mediated proteolysis</td>
<td>Cdc34, Ube2g1, Cud5, Frx1, Ube2g2</td>
<td>0.094154</td>
</tr>
<tr>
<td>mmu04910:Insulin signaling pathway</td>
<td>Calm1, Ppmp1, Pppl1cc, Foxos1, Nras</td>
<td>0.098052</td>
</tr>
</tbody>
</table>

Based on the above analysis results, it can be inferred that the period from the injection of OA to 1.5 h is the initiation stage of inflammation. As more and more pathways become involved, such as the B cell, T cell, and natural killer cell pathways, the inflammatory...
response proceeds and results in more abnormal reactions, which ultimately cause substantial lung injury.

**DISCUSSION**

In this study, we explored the genes and biological pathways implicated in ALI using a mouse model administered with OA. Through analyzing the differential co-expression patterns, we found that most of the DCGs were identified at 1.5 and 3 h. A biological interpretation of the DCGs was performed, which revealed several enriched pathways. Ubiquitin-mediated proteolysis, hematopoietic cell lineage, leukocyte transendothelial migration, and spliceosomes were enriched at 1.5 h. The B cell receptor signaling pathway, T cell receptor signaling pathway, natural killer cell-mediated cytotoxicity, and the Fc epsilon RI signaling pathway were significantly enriched at 3 h.

Lesur et al. (2010) conducted a histological examination for the ALI mouse model and reported that lung injury proceeds 1.5 h after OA injection along with interstitial neutrophil accumulation and edema; however, the lesions halted and edema disappeared 3 h after injection. We reached similar conclusions according to the differential co-expression patterns. We speculated that there was a peak between 1.5 and 3 h, which showed the maximum number of DCGs. As the stimulus gradually disappeared over time, the tissue restored to the normal state, and thus few DCGs were identified at 4 h and later. These results suggested that the DCGs could play important roles in the initiation and development of ALI.

Functional analysis for the DCGs revealed that ubiquitin-mediated proteolysis was the most significantly enriched pathway. Previous studies have indicated its regulatory role in the immune system response (Natoli and Chiocca, 2008; El-Hashim et al., 2011). Ubiquitination is involved in the development of the immune system and in regular immune responses through affecting the stability or activity of proteins. For example, ubiquitination is closely associated with the activation of nuclear factor-kappa B, which is an important transcription factor and controls the immune response, inflammatory response, and apoptosis process (Chen, 2005; Bhoj and Chen, 2009; Sun, 2012). E3 ubiquitin protein ligases play a big role in the regulation process, and some of them have also been shown to be involved in the inflammatory response, including CBLB (Bachmaier et al., 2007). Murine double-minute 2 (MDM2) and suppressor of cytokine signaling 3 (SOCS3), which were identified in our study, are also associated with inflammation. MDM2 is believed to link inflammation with epithelial healing during acute kidney injury (Mulay et al., 2012), while SOCS3 is able to attenuate pro-inflammatory signaling (Jo et al., 2005; White et al., 2011). Ligases, such as CUL3, CUL5, and CUL7 were also found to be differentially co-expressed. Studies have revealed their association with cancer (Kim et al., 2007; Lubbers et al., 2011; Thu et al., 2011) and embryonic development (Kasper et al., 2006), but their roles in ALI remain to be elucidated. Furthermore, studies have shown that a variety of ubiquitin ligases are involved in preventing the immune system from attacking self tissues, and that the dysfunction of ubiquitin ligase could lead to the development of autoimmune diseases (Deng et al., 2000; Uchida et al., 2004; Bottomley et al., 2005; Bhoj and Chen, 2009).

The hematopoietic cell lineage and leukocyte transendothelial migration were also significantly enriched at 1.5 h. The hematopoietic cell lineage is mainly involved in the development of blood cells, including granulocytes, monocytes, and dendritic cells. Leukocyte
transendothelial migration is also associated with the inflammatory response. When acute inflammation occurs, local tissues produce interleukin (IL)-1β, IL-8, tumor necrosis factor (TNF)-α, and other inflammatory cytokines to up-regulate the adhesion molecules of vascular endothelial cells. Meanwhile, neutrophils adhere to the surface of injured endothelial cells and then transfer into the interstitial and alveolar spaces, thereby releasing a large number of pro-inflammatory mediators such as inflammatory cytokines, peroxides, leukotriene, protease, and platelet-activating factor, and is thus involved in the neutrophil-mediated inflammatory response and lung injury (Dellinger et al., 2008). Migration of leukocytes from the blood into the tissues is important for immunological surveillance and the inflammatory processes. However, if the leukocytes fail to handle the damaged tissues, they will prolong the process and pose damage to self tissues. Indeed, histological observations have shown neutrophil accumulation in the ALI mouse model (Lesur et al., 2010); therefore, it is not surprising that these two pathways were enriched at corresponding time points. A literature search was performed for some of the DCGs involved in these pathways to further understand their roles in the immune response and in ALI. Previous studies have demonstrated that Cd5 and Csf1 are associated with the immune system (Mizoguchi et al., 2003; Loureiro et al., 2008). The protein product of Vcam1 mediates leukocyte-endothelial cell adhesion and signal transduction (Lobb et al., 1995). Ezrin and Vinculin are also known to play critical roles in cell adhesion, migration, and organization (Hiscox and Jiang, 1999; Poullet et al., 2001; Carisey and Ballestrem, 2011; Shen et al., 2011).

Another unexpected pathway is the spliceosome pathway. The initial pre-mRNA transcription product must go through the process of splicing to form mature mRNA. As many pro- and anti-inflammatory cytokines are produced, more and more spliceosomes are required. Therefore, alternative splicing might also contribute to the change in expression level. Alternative splicing produces a high degree of protein diversity under conditions of low genetic value, which is valuable for immune systems. The DCGs Prpf6, Sf3a2, and Phf5a encode subunits of the spliceosome pathway.

CONCLUSIONS

Overall, we identified hundreds of DCGs and confirmed their associations with inflammation as well as ALI through functional analysis. These DCGs are potential drug targets, but require further validation to confirm their usefulness.

ACKNOWLEDGMENTS

Research supported by the Shanghai Medical Key Subject Construction Project (#ZK2012A28), the National Clinical Key Specialty Construction Project and the Pudong New Area of Outstanding Young Medical Talents (#PWRq2012-13).

REFERENCES
