Analysis of Transcriptional Profiles to Discover Biomarker Candidates in *Mycobacterium avium* subsp. *paratuberculosis*-Infected Macrophages, RAW 264.7

Seung Bin Cha†, Anna Yoo†, Hong Tae Park, Kyoung Yong Sung, Min Kyoung Shin, and Han Sang Yoo*

Department of Infectious Diseases, College of Veterinary Medicine, Brain Korea 21 for Veterinary Science, Seoul National University, Seoul 151-742, Republic of Korea

Paratuberculosis (PTB) or Johne’s disease (JD) is one of the most serious chronic debilitating diseases of ruminants worldwide that is caused by *Mycobacterium avium* subsp. *paratuberculosis* (MAP). MAP is a slow-growing bacterium that has very long latent periods, resulting in difficulties in diagnosing and controlling the disease, especially regarding the diagnosis of fecal shedders of MAP without any clinical signs. Based on this situation, attempts were made to identify biomarkers that show early responses to MAP infection in a macrophage cell line, RAW 264.7. In response to the infection with the bacterium, a lot of genes were turned on and/or off in the cells. Of the altered genes, three different categories were identified based on the time-dependent gene expression patterns. Those genes were considered as possible candidates for biomarkers of MAP infection after confirmation by quantitative RT-PCR analysis. To the best of our knowledge, this is the first attempt at discovering the host transcriptomic biomarkers of PTB, although further investigation will be required to determine whether these biomarker candidates are associated within the natural host.

**Keywords:** Johne’s disease, biomarker, transcriptional change

Introduction

Paratuberculosis (PTB) or Johne’s disease (JD) is one of the most serious chronic debilitating diseases of ruminants worldwide that is caused by *Mycobacterium avium* subsp. *paratuberculosis* (MAP). The disease has been seen primarily in cattle, sheep, and goats; however, recently, a possible association between MAP and Crohn’s disease (CD) in humans has been focused on owing to similarities in these chronic inflammatory bowel diseases and the evidence supporting this linkage [1, 18, 20, 21].

MAP is a slow-growing, acid-fast intracellular bacillus that has very long latent periods, resulting in difficulties in diagnosing and controlling the disease. PTB typically progresses through subclinical and clinical stages [33]. Because cattle in the subclinical stage may shed the bacteria, the first step in the eradication of PTB is the development of diagnostic tests that can identify and remove potential fecal shedders of MAP during the early stage, thereby preventing the spread of JD [8, 31]. However, current diagnostic tests, including the most widely used serological tests and direct organism tests, have limitations, especially in terms of the diagnosis of fecal shedders of MAP without any clinical signs [31, 33].

Biomarkers provide prognostic information about future health status, as well as an advanced knowledge of pathogenesis [29]. Recently, biomarkers for early diagnosis have been studied in many cases, especially in tuberculosis [3, 9, 11, 19]. However, not much research is ongoing regarding biomarkers in PTB, except for one study involving serum proteomic profiling [33]. Considering the fact that the replication of MAP occurs primarily inside the macrophage of the infected host [6], we infected the professional phagocyte RAW 264.7 with MAP for 6, 24, 48, and 120 h. The
transcriptional changes in these macrophages were compared to discover candidates for biomarkers via the analysis of transcriptomes.

Materials and Methods

Bacterial Infection and RNA Preparation

The bacterial strain Mycobacterium avium subsp. paratuberculosis ATCC 19698 was grown on 7H10 agar containing oleic acid-albumin-dextrose-catalase (OADC; Difco Laboratories, USA), mycobactin J (IDEXX), casitone (BD, USA), glycerol, and Bactec Meit supplement (BD, USA) at 37°C for 3–4 weeks. The fresh culture was suspended in PBS and diluted to an optical density (OD) of 1.0 for 1 × 10^7 cells/ml [23]. A mouse leukemic monocyte macrophage cell line, RAW 264.7, was cultivated in complete Modified Eagle’s minimal essential medium (DMEM; Invitrogen, USA) supplemented with 10% fetal bovine serum (FBS; Invitrogen, USA) and Anti-Anti (Antibiotic-Antimycotic; Invitrogen, USA) at 37°C under a 5% CO₂ atmosphere. For bacterial infection, we followed the protocols described previously [10]. Briefly, 2 × 10^6 cells/well were allowed to adhere to a 6-well tissue culture plate with DMEM and without antibiotics, and the prepared bacterial cells were inoculated. MAP ATCC 19698 was used at a multiplicity of infection (MOI) of 20:1. Cells were washed twice with fresh media after 2 h of infection in order to wash off non-adherent bacteria and were incubated with fresh media. RAW 264.7 cells were collected to extract the total RNA at 6, 24, 48, and 120 h post-infection (pi). The total RNA was isolated from infected RAW 264.7 cells using an RNeasy Mini Kit (Qiagen, USA). This experiment was conducted in duplicate.

Microarray

The total RNA was amplified and purified using the Ambion Illumina RNA amplification kit (Ambion, USA) in order to yield biotinylated cRNA according to the manufacturer’s instructions. First, 1.5 µg of labeled cRNA samples was hybridized to each mouse-6 expression bead array for 16–18 h at 58°C according to the manufacturer’s instructions (Illumina, Inc., USA). An Illumina expression beadchip (MouseWG-6 v2.0) that can cover more than 45,000 transcripts was used in this study. The detection of the array signal was carried out using Amersham fluorolink streptavidin-Cy3 (GE Healthcare Bio-Sciences, UK) following the bead array manual. The arrays were scanned with an Illumina bead array reader confocal scanner, according to the manufacturer’s instructions. The quality of hybridization and overall chip performance were monitored via visual inspection, using both internal quality control checks and the raw scanned data. The raw data were extracted using the software provided by the manufacturer (Illumina GenomeStudio ver. 2011.1 (Gene Expression Module ver. 1.9.0)). RNA amplification, labeling, array hybridization, and scanning were carried out by Macrogen Inc. (Republic of Korea).

Statistical Analysis

Array probes that have detection p-values of 0.05 (similar to signal to noise) in over 50% of samples were filtered out. (We applied a filtering criterion for data analysis; a higher signal value was required to obtain a detection p-value < 0.05.) The selected gene signal value was transformed via its logarithm and normalized by using the quantile method. The statistical significance of the expression data was determined using fold change and an LPE test in which the null hypothesis was that no difference exists between the two groups. The false discovery rate (FDR) was controlled by adjusting the p-value using the Benjamini-Hochberg algorithm. Hierarchical cluster analysis was performed using complete linkage and Euclidean distance as a measure of similarity. Gene-enrichment and functional annotation analysis for the significant probe list was performed using DAVID (http://david.abcc.ncifcrf.gov). All data analysis and visualization of the variously expressed genes were conducted using R 2.15.1 (http://www.r-project.org).

Validation of Microarray Results

To validate the microarray results, five altered genes from the 6 h pi samples and another five altered genes from the 120 h pi samples (Table 1) were randomly selected and submitted to quantitative RT-PCR. The total RNA from the macrophage, the remainder after microarray analysis, was used for reverse transcription with a SuperScript VILO cDNA synthesis kit (Invitrogen, USA) according to the manufacturer’s protocol. The RT-PCR was performed with 1 µl of cDNA using a Rotor-Gene SYBR Green PCR kit and a Rotor-Gene Q real-time PCR cycler (Qiagen). Amplification was performed for 35 cycles at 95°C for 15 sec, followed by 45 sec at 60°C, detecting fluorescence during the extension phase. The expression level was determined via the 2^-ΔΔCt method [16], using a housekeeping gene, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), as a reference. The relative expression level was compared with that of an uninfected macrophage, used as a control, in order to determine the expression-fold change of each gene.

Results

Microarray Analysis of Differentially Expressed Genes Following MAP Infection

This study used microarrays to discover biomarker candidates using the Illumina Mouse WG-6 v2 Expression BeadChip, which covers more than 45,000 transcripts. The threshold value for the microarray was chosen with an expression change of ≥2-fold for both up- and down-regulated genes with a p-value of less than 0.05. The numbers of genes that satisfied this criterion were shown (Fig. 1), and the expression levels were displayed via scattered plots, with the median being the normalized hybridization signals of macrophages (Fig. 2). They showed many differences in gene expression level in accordance with time. Two hundred and twenty-four (0.73%) and 261
(0.85%) genes were up-regulated and 184 (0.6%) and 241 (0.78%) genes were down-regulated in macrophages at 6 and 120 h pi, respectively, as compared with uninfected macrophages. However, the expression levels of much lower numbers of genes (75 genes up-regulated and 147 genes down-regulated) were altered at 24 h pi, and only 22 genes were altered at 48 h pi.

**Comparisons of Gene Expression Profiles of Macrophages at 6 and 120 h pi**

Based on the criterion that biomarkers for early diagnosis should represent a specific feature during the early phase of infection, we focused on the earliest and the very last phases of infection in this study. Interestingly, 570 (1.85%) genes were up-regulated and 593 (1.92%) genes were down-regulated when we compared 6h pi macrophages to 120h pi macrophages (Fig. 1). A full list of genes with altered expression levels is provided in Supplementary Table S1.

**Identification of Differently Expressed Genes at 6, 48, and 120 h pi**

In the results of the hierarchical clustering analysis, which was intended to find similar patterns in the expression level of each gene, it can be divided into three groups (Fig. 3). Among the 1,163 genes expressed differently at 6 and 120 h pi, we selected genes with characteristic features during the experiment. Three genes (Tfrc, Cx3cr1, and Ccne2) with increased expression at 6 h pi and decreased expression at 120 h pi were classified as Group A. Fourteen genes (Cox6a2, Gdf15, Ypel3, Aqp9, Slc40a1, Tmem154, Cd74, Aatk, Rras, Gadd45a, Ypel5, Hebp1, Ens2, and Macrod1) with decreased expression at 6 h pi and increased expression at 120 h pi were classified as Group B. The remaining genes were classified into Group C.

### Table 1. Primers used for qRT-PCR.

<table>
<thead>
<tr>
<th>Accession No.</th>
<th>Gene symbol (Description)</th>
<th>Forward primers (5’ → 3’)</th>
<th>Reverse primers (5’ → 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM_013693.2</td>
<td>Tnf (tumor necrosis factor)</td>
<td>CAACGCCCTCTGGGCAACG</td>
<td>TCGGCGCGCTGTCCCTTT</td>
</tr>
<tr>
<td>NM_008392.1</td>
<td>Irg1 (immunoresponsive gene 1)</td>
<td>CCGTGCCTCTGTCGTGCAG</td>
<td>CGTGCACCAGCTTGCCGTTG</td>
</tr>
<tr>
<td>NM_011683.4</td>
<td>Tfrc (transferrin receptor)</td>
<td>AGTAGCCTCTGGAGTGTG</td>
<td>TCTGTACCAGCTGCAATC</td>
</tr>
<tr>
<td>NM_009943.2</td>
<td>Cox6a2 (cytochrome c oxidase subunit Vla polypeptide 2)</td>
<td>AGCCCTCTGCTCCCTTAACT</td>
<td>CGGATCCGAGTTGTCGATAC</td>
</tr>
<tr>
<td>NM_025347.1</td>
<td>Ypel3 (Yippee-like 3)</td>
<td>CTGTGCCTGACATCCACTCTGT</td>
<td>TGTCCCTCAAAGGCTGCTCA</td>
</tr>
<tr>
<td>NM_009987.2</td>
<td>Cx3cr1 (chemokine (C-X3-C) receptor 1)</td>
<td>ACAAGAGACTCCACGCTGAC</td>
<td>GTGCAAGCAACAGAGTTG</td>
</tr>
<tr>
<td>NM_008361.1</td>
<td>Gadd45a (growth arrest and DNA-damage-inducible 45 alpha)</td>
<td>AGACCGAAAGGATGACCG</td>
<td>GGTCTACGTGGCAGCAGT</td>
</tr>
<tr>
<td>NM_009635.2</td>
<td>Arvl (advillin)</td>
<td>AGGACTCTCCCCAAGACGA</td>
<td>CTGCCCTAAGTCGCCACGGA</td>
</tr>
<tr>
<td>NM_008491.1</td>
<td>Lcn2 (lipocalin 2)</td>
<td>GCCCAATCGACTCTGGGAAA</td>
<td>TGGCGAACCTGTTGAGTC</td>
</tr>
<tr>
<td>NM_013538.4</td>
<td>Cdc43 (cell division cycle-associated 3)</td>
<td>CTACTGCAGTTGTCGGCGC</td>
<td>AACCCATCTCCACTGGGCC</td>
</tr>
</tbody>
</table>

**Fig. 1.** Numbers of genes with altered expression levels during this experiment ($p < 0.05$, |fold change| ≥ 2). RNA was isolated from RAW 264.7 macrophages at 6, 24, 48, and 120 h pi (post-infection).
Fig. 2. Plots of the expression levels compared at each time point. Red dots indicate an expression level change of $|2|^\pm$-fold for both up- and down-regulated genes. The expression level was calculated by using a base-2 logarithm of the normalized hybridization signals from each sample.
120 h pi were classified as Group B. However, there were no genes with constantly increased or decreased expression during the 120 h of infection. For this reason, to cover the mid-phase (48 h pi) of the infection period, three genes (Irf7, Nfkbiz, and Lcn2) that were mostly increased at 48 h pi were selected and classified as Group C. The selected genes and their descriptions, as well as the expression kinetics of the selected genes, are shown in Table 2 and Fig. 4, respectively.

Validation of Microarray Data
To confirm the microarray data, we performed quantitative RT-PCR with 10 genes that showed altered expression levels. We selected the Tnf, Irg1, Tfrc, Cox6a2, and Ypel3 genes from the 6 h pi group and the Cx3cr1, Gadd45a, Avil, Lcn2, and Cdca3 genes from the down-regulated gene group. We could validate the microarray data because all genes tested via qRT-PCR showed greater fluctuations (increases or decreases) than those found in the microarray data, but the fluctuations were in the same direction (Fig. 5).

Discussion
PTB has a prolonged subclinical stage that can shed bacteria without any clinical signs. Because these fecal shedders might act as sources of infection to other animals, the development of a diagnostic method that is useful in the early stage is very important to eradicate or control the disease. In light of this, the aim of this study was to
discover biomarker candidates from the transcriptome of the macrophage, in which MAP may persistently survive and replicate.

Although several studies have already analyzed the gene expression profiles of peripheral blood mononuclear cells (PBMC) or infected tissues in bovines [2, 22], which are a natural host of MAP, we analyzed the gene expression profiles in MAP-infected murine macrophages, which have a much wider variety available in terms of selecting biomarker candidates. Similar expression patterns have been observed in our results as in previous work with bovine PBMC [2]: a great number of genes were activated during the early phase of in vitro infection (6 h pi) and returned to near or below their original levels at the mid phase of in vitro infection (48 h pi). However, this phenomenon was delayed for several hours as compared with the data from the bovine PBMC, suggesting the host preference of this bacterium. However, several genes began to show

![Hierarchical clustering analysis of the expression levels of each gene.](image)

**Fig. 3.** Hierarchical clustering analysis of the expression levels of each gene.

![Fold changes of selected genes in each group from RAW 264.7 cells exposed to Mycobacterium avium subsp. paratuberculosis strains over time.](image)

**Fig. 4.** Fold changes of selected genes in each group from RAW 264.7 cells exposed to Mycobacterium avium subsp. paratuberculosis strains over time.

In each panel, the period of infection with MAP is depicted along the x-axis, and the mean expression difference, represented as the fold change relative to uninfected cells, is depicted along the y-axis. (For group B, three genes with the most remarkable change were selected and depicted.)
altered expression levels during the late phase of in vitro infection (120 h pi). Moreover, similar functional changes occurred in this experiment (data not shown) as in a recent study with bovine mononuclear-derived macrophages (MDMs) [10], suggesting that the in vitro murine model could reflect the pathogenesis in bovines, which are a natural host of PTB.

As the aim of our study was to discover biomarker candidates, we approached the altered gene during each pi period, rather than focusing on a functional discussion of pathogenesis that had already been studied in natural hosts [2, 8, 17]. Based on a previous study with natural hosts [10] that found that macrophage genes were differently expressed during the time of infection according to the Cluster of Orthologous Genes (COG) classification system and that the most variable gene alteration occurred at 6 h pi as compared with 120 h pi, we selected the biomarker candidates from the genes that were up-regulated at 6h pi but down-regulated at 120 h pi, and vice versa, and designated these to be Groups A and B, respectively.

The three genes in Group A were Tfrc, Cx3cr1, and Ccne2. Tfrc, transferrin receptor, is known to be expressed in many tumor types and has recently been suggested to be a common cancer biomarker [4]. Cx3cr1 and Ccne2 were revealed as a biomarker candidate for obesity [25] and a prognostic marker for lymph node-negative breast cancer patients [24], respectively. Additionally, in Group B, most genes had been identified as biomarker candidates or prognostic markers for many diseases. Gdf 15 and Agp9 were found in cardiovascular disease patients [13, 28], and Cd74, Gadd45a, and Eno2 were found in various types of cancers [5, 26, 30]. Ypel3, Tmem154, and Rras were found to induce cellular senescence [12], to be involved in lentivirus infection in sheep [7], and to be an important biomarker in biliary atresia [32], respectively. Moreover, other genes in this group have been revealed to be involved in many cases, especially in cancer research or multiplexed assays for a broad range of biomarker studies. As biomarkers can perform substantially better than any other markers when combined [29] and as there was no gene that was steadily up- or down-regulated during this experiment, we suggested three of the most increased genes at 48 h pi (Group C). Irf7, Nfkbia, and Lcn2 were found to be a potential biomarker for the successful selective killing of cancer cells via oncolytic

### Table 2. Selected genes increased at 6 h pi but decreased at 120 h pi, or vice versa (p < 0.05, |fold change | ≥ 2).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession no.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>Tfrc</td>
<td>NM_011638.3 Transferrin receptor</td>
</tr>
<tr>
<td></td>
<td>Cx3cr1</td>
<td>NM_009987.2 Chemokine (C-X3-C) receptor 1</td>
</tr>
<tr>
<td></td>
<td>Ccne2</td>
<td>NM_001037134.1 Cyclin E2, transcript variant 1</td>
</tr>
<tr>
<td>Group B</td>
<td>Cox6a2</td>
<td>NM_009943.2 Cytochrome c oxidase subunit V1a polypeptide 2</td>
</tr>
<tr>
<td></td>
<td>Gdf15</td>
<td>NM_011819.1 Growth differentiation factor 15</td>
</tr>
<tr>
<td></td>
<td>Ypel3</td>
<td>NM_025347.1 Yippee-like 3 (Drosophila)</td>
</tr>
<tr>
<td></td>
<td>Agp9</td>
<td>NM_022026.2 Aquaporin 9</td>
</tr>
<tr>
<td></td>
<td>Slc40a1</td>
<td>NM_016917.2 Solute carrier family 40 (iron-regulated transporter), member 1</td>
</tr>
<tr>
<td></td>
<td>Tmem154</td>
<td>NM_177260.2 Transmembrane protein 154</td>
</tr>
<tr>
<td></td>
<td>Cd74</td>
<td>NM_001042605.1 CD74 antigen (invariant polypeptide of major histocompatibility complex, class II antigen-associated)</td>
</tr>
<tr>
<td></td>
<td>Aatk</td>
<td>NM_007377.3 Apoptosis-associated tyrosine kinase</td>
</tr>
<tr>
<td></td>
<td>Ras</td>
<td>NM_009101.2 Harvey rat sarcoma oncogene, subgroup R</td>
</tr>
<tr>
<td></td>
<td>Gadd45a</td>
<td>NM_007836.1 Growth arrest and DNA-damage-inducible 45 alpha</td>
</tr>
<tr>
<td></td>
<td>Ypel5</td>
<td>NM_027166.5 Yippee-like 5 (Drosophila)</td>
</tr>
<tr>
<td></td>
<td>Hebp1</td>
<td>NM_013546.2 Heme-binding protein 1</td>
</tr>
<tr>
<td></td>
<td>Eno2</td>
<td>NM_013509.2 Enolase 2, gamma neuronal</td>
</tr>
<tr>
<td></td>
<td>Macrod1</td>
<td>NM_134147.3 MACRO domain-containing 1</td>
</tr>
<tr>
<td>Group C</td>
<td>Irf7</td>
<td>NM_016850.2 Interferon regulatory factor 7</td>
</tr>
<tr>
<td></td>
<td>Nfkbia</td>
<td>NM_030612.2 Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta</td>
</tr>
<tr>
<td></td>
<td>Lcn2</td>
<td>NM_008491.1 Lipocalin 2</td>
</tr>
</tbody>
</table>
viral therapy [15], to be differentially expressed in local primary tumors as compared with metastatic primary tumors [27], and to be a biomarker predicting the development of delayed graft function [14], respectively. Therefore, the biomarker candidates for PTB found in this study shared common features with the biomarker candidates for other diseases and cancers. Considering the fact that a good biomarker is one that is specific, robust, and applicable to clinical usage [22], our data may provide basic information for biomarker discovery for other diseases.

In summary, we identified several biomarker candidates for the early diagnosis of PTB by analyzing in vitro transcriptional changes in host cellular responses during the progression of the infection. To the best of our knowledge, this is the first attempt at discovering host transcriptomic biomarkers with PTB, although it will require further investigation to determine whether these biomarker candidates are associated within the natural host or not.

**Fig. 5.** Validation of microarray data via quantitative RT-PCR. The relative expression level was normalized in terms of the GAPDH expression level relative to uninfected cells via the 2^(-∆∆CT) method.

**Acknowledgments**

This study was supported by the Rural Development Administration (PJ008970012012), BK21 for Veterinary Science, and the Research Institute for Veterinary Science, Seoul National University, South Korea.

**References**