# Establishment of combined analytical method to extract the genes of interest from transcriptome data

Biochemistry and Biophysics Reports 7 (2016) 63–69

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

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

# Establishment of combined analytical method to extract the genes of interest from transcriptome data

Akinori Ishihara [a,b], Hidenori Yamanaka [c], Reiko Takahashi [d], Tomomi Numajiri [d], Saki Kaneko [d], Yoko Ishizawa [d], Sakuji Koya [d], Kiyoshi Yamauchi [a,b,*]

[a] Department of Biological Science, Faculty of Science, Shizuoka University, 836 Ohya, Shizuoka 422-8529, Japan
[b] Green Biology Research Division, Research Institute of Green Science and Technology, Shizuoka University, 836 Ohya, Shizuoka 422-8529, Japan
[c] Chemicals Assessment Center, Chemicals Evaluation and Research Institute, 1600, Shimo-Takano, Sugitomachi, Kitakatsushika-gun, Saitama 345-0043, Japan
[d] Wakanyaku Medical Institute, Ltd., 1193 Fujimimachi Akagiyama, Maebashi 371-0101, Japan

ARTICLE INFO

*Article history:*
Received 7 December 2015
Received in revised form
23 April 2016
Accepted 18 May 2016
Available online 19 May 2016

*Keywords:*
Proteome
Transcriptome
Crude drug
Pathway

ABSTRACT

Techniques for analyzing genome-wide expression profiles, such as the microarray technique and next-generation sequencers, have been developed. While these techniques can provide a lot of information about gene expression, selection of genes of interest is complicated because of excessive gene expression data. Thus, many researchers use statistical methods or fold change as screening tools for finding gene sets whose expression is altered between groups, which may result in the loss of important information. In the present study, we aimed to establish a combined method for selecting genes of interest with a small magnitude of alteration in gene expression by coupling with proteome analysis. We used hypercholesterolemic rats to examine the effects of a crude herbal drug on gene expression and proteome profiles. We could not select genes of interest by using standard methods. However, by coupling with proteome analysis, we found several effects of the crude herbal drug on gene expression. Our results suggest that this method would be useful in selecting gene sets with expressions that do not show a large magnitude of alteration.



## 1. Introduction

Recently, methods for analyzing genome-wide expression profiles, such as microarrays and next-generation sequencing, have been developed. These methods enable us to obtain a large amount of gene expression data, easily and inexpensively. While the massive amount of data provides an overview of certain biological phenomena on the basis of suitable statistical analyses, there are several limitations. In general, analysis of genome-wide expression data is carried out by: (1) normalization of the expression data, (2) selection of genes of interest from the data, and (3) visualization of expression profiles [1]. The most important step is the selection of genes of interest and determining the biological significance of those selected gene set. In addition, most researchers select genes with statistically altered expression or when the magnitude of the fold change between groups is large, because it is difficult to limit the numbers of genes selected by using small thresholds [2–4]. However, such methods for the selection of genes often result in the loss of information because genes that demonstrate large variation in expression within groups or minor changes in expression between groups are not selected as genes of interest. Two-dimensional difference gel electrophoresis (2D-DIGE) is a new technique that compares the abundance of a large number of proteins, and many researchers have attempted to use this technique in various areas of research [5,6]. However, compared to studies utilizing model organisms that have sufficient genomic information, there are limitations in analyzing the proteome of non-model organisms whose genomes have not been sequenced [7], and the throughput of the experiment is lower than that of transcriptome analysis. Additionally, both proteome and transcriptome analyses have a similar problem in the selection of candidates, that is, loss of information. Thus, transcriptome and proteome analyses have both merits and demerits. It is necessary to develop new methods for analyzing transcriptome and proteome data without the loss of information.

In our previous study, we used microarray analysis to determine the effects of a combined extract of *Sasa albo-marginata* leaves (kumazasa; *Sasa kurinensis* Makino), Japanese red pine leaves (*Pinus densiflora* Sieb. et Zucc), and ginseng roots (*Panax*

* Corresponding author at: Green Biology Research Division, Research Institute of Green Science and Technology, Shizuoka University, Ohya 836 Suruga-ku, Shizuoka 422-8529, Japan.
*E-mail address:* yamauchi.kiyoshi@shizuoka.ac.jp (K. Yamauchi).

http://dx.doi.org/10.1016/j.bbrep.2016.05.015

*ginseng* C.A. Meyer) (SJG) on high-cholesterol diet (HCD)-induced hypercholesterolemia [8]. In that study, we finally confirmed that expression of the gene encoding cytochrome P450 7A1 was altered in response to HCD and HCD with SJG, although we could not select the gene as a candidate gene of interest by using standard microarray analysis because of the minor change in gene expression [8]. This result suggests that it may be difficult to determine the molecular mechanism underlying the action of crude drugs using standard methods because changes in gene or protein expression caused by crude herbal drugs may be smaller than those caused by most pharmaceutical drugs. However, several plant polyphenols are known to play a major role in reducing risk of arteriosclerosis [9–13], suggesting that small molecules from plants have useful health effects. In this study, we aimed to select genes of interest whose transcript or protein expression profiles showed small changes by using transcriptome analysis coupled with proteome analysis and used crude herbal drug-treated rats as model animals.

## 2. Materials and methods

### 2.1. Reagents

Immobilized pH gradient strip (pH 3–10), pharmalytes, drystrip cover fluid, bromophenol blue, agarose, Cy2, Cy3, and Cy5 were purchased from GE Healthcare (Little Chalfont, UK). Protease inhibitor cocktail (Pefabloc SC and Pefabloc SC protector) was purchased from Roche (Mannheim, Germany). Sequence-grade trypsin was obtained from Promega UK (Southampton, Hants, UK). All other chemicals used in this study were of the highest grade available and were purchased from GE Healthcare, Dojin Chemical Japan (Osaka, Japan), Sigma (St. Louis, MO), Wako (Osaka, Japan), Nacalai Tesque (Kyoto, Japan), or Kanto (Tokyo, Japan). The crude herbal drug SJG was prepared by Wakanyaku Medical Institute, Ltd. (Maebashi, Japan). SJG is composed of a water extract of kumazasa leaves and ethanol extracts of Japanese red pine leaves and ginseng roots in the ratio 8:1:1 [14]. SJG was supplied as a liquid preparation and diluted using tap water to 50% (v/v). We previously analyzed the component of SJG and found that SJG contained several compounds like tricin, *p*-coumaric acid, ginsenoside [8].

### 2.2. Animals and treatment

Twenty male Wistar rats (age, 13 weeks) were placed in a room with controlled lighting (12-h light/dark cycle), temperature ($23 \pm 1$ °C), and humidity ($55 \pm 10$%). All rats were provided with food and tap water or 50% SJG ad libitum throughout the experiment. We divided the rats into four groups (Groups 1, 2, 3, and 4). After 1 week of habituation, all rats were fed a standard diet (23% protein, 5% fat, and 55% nitrogen-free extract; MF, Oriental Yeast Co., Tokyo, Japan) and given water (Groups 1 and 3) or 50% SJG (Groups 2 and 4) for 12, 30, or 52 weeks (experimental phase I, Fig. 1). The rats were then provided with food and water as follows: standard diet and water, Group 1; standard diet and 50% SJG, Group 2; HCD and water, Group 3; and HCD and 50% SJG, Group 4 for three weeks (experimental phase II, Fig. 1). HCD was the standard diet supplemented with 0.5% cholesterol and 0.5% cholic acid. Then, the rats were anesthetized with diethyl ether and sacrificed. The plasma was collected for measuring the levels of biochemical parameters, such as glucose, total cholesterol, high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, cholesteryl ester, free cholesterol, triglyceride (TG), free fatty acid (FFA), phospholipids, and total lipids, and the livers were collected for RNA extraction. The animal experiments were approved by the Animal Research Committee of Wakanyaku Medical Institute, Ltd., and performed in accordance with the Guidelines for Care and Use of Laboratory Animals at Wakanyaku Medical Institute, Ltd. and the Guidelines for Proper Conduct of Animal Experiments from Science Council of Japan.

### 2.3. Plasma biochemical parameters

Blood samples were obtained from the postcaval vein by using a vacuum blood collection tube containing heparin. Plasma levels of glucose, total cholesterol, HDL cholesterol, LDL cholesterol, cholesteryl ester, free cholesterol, TG, FFA, phospholipids, and total lipids were determined using the Glucose test Wako (Wako), L-type Wako cholesterol (Wako), Cholestest N HDL (Sekisui Medical Co., Ltd, Tokyo, Japan), Cholestest LDL (Sekisui), L-type Wako cholesterol (Wako) and L-type Wako free cholesterol (Wako), L-type Wako free cholesterol (Wako), L-type Wako TG (Wako), NEFA-SS Eiken (Eiken Chemical Co., Ltd, Tokyo, Japan), L-type Wako phospholipids (Wako), and Total lipids reagent Kokusai (Sysmex, Kobe, Japan), respectively.

### 2.4. RNA isolation and microarray analysis

Isolation of total RNA from the liver and microarray analyses were performed according to the methods described in our previous study [8]. Total RNA was extracted from frozen rat livers by using the High Pure RNA Isolation Kit (Roche) and subjected to oligonucleotide array analysis (Agilent Microarray Design ID 028279; Agilent, Santa Clara, CA, USA) by using a single-color labeling system (Low Input Quick Amp Labeling Kit for One Color; Agilent). We performed *t* tests or Benjamini and Hochberg's method to compare signals between two groups, and a Venn diagram was created to extract genes whose expression was altered by the intake of HCD, and the expression was further modified by SJG. All analyses for microarray data were performed using linear models for microarray data (limma) package [15] or Subio


**Fig. 1.** Scheme of the experimental schedule. All rats were acclimatized for a week and fed a standard diet throughout habituation and the experimental phase I (12, 30, and 52 weeks). After habituation, the rats were divided into four groups. Groups 1 and 3 received water throughout the experimental phase, while groups 2 and 4 received 50% SJG. In experimental phase II (3 weeks), groups 3 and 4 were fed a high-cholesterol diet. Finally, all the rats were sacrificed, and plasma and liver were collected for analysis.

platform software, and we finally identified cholesterol-regulated genes whose expressions were altered by SJG (Supplementary Fig. S1A and B). To select the genes whose expression patterns were similar to certain gene, first of all, we obtained average vector of the gene. After that we calculated the Peason's correlation between the gene and remaining all genes on one-on-one level. Finally, we selected the genes which had high correlation ($r > 0.85$).

### 2.5. 2D-DIGE

2D-DIGE was performed as described by Yamanaka et al. [16]. In brief, liver samples were homogenized in 10 volumes of lysis buffer that contained 4% (w/v) CHAPS, 2 M thiourea, 8 M urea, 10 mM Tris–HCl (pH 8.8), 4 mM Pefabloc SC, 20 mM Pefabloc SC protector on ice by using a Potter type homogenizer, followed by sonication. After centrifugation was performed at 20,800 × *g* for 20 min at 10 °C, the supernatant was collected as the protein sample. Protein concentration of each sample was determined using the dye-binding method [17] with transferrin as the standard, and protein samples (50 μg) were labeled with 200 pmol of *N*-hydroxy succinimidyl ester-derivatives of cyanine dyes Cy3 or Cy5. A mixture of all protein samples was labeled with Cy2 and used as the standard on all gels. Labeled samples and the standard were run on the same gel, and 2-dimensional electrophoresis was performed. We used wide-range immobilized pH gradient strips (pH 3–10, 24 cm). Separation in the first dimension was performed using Multiphor II (GE Healthcare) at 54 kV h and 20 °C under darkness. Then, we used 12% T, 7.5% C gels (24 cm × 20 cm) for separation in the second dimension with Ettan DALT twelve (GE Healthcare) at 2 W/gel and 20 °C. The gels were scanned directly with the 2920 2D-Master Imager (GE Healthcare), and the scanned images were analyzed using the Decyder (Version 4.00) software (GE Healthcare) to quantify the amount of proteins. We performed *t* tests to compare signals between two groups, and a Venn diagram was created to select the proteins whose levels were altered by HCD intake, and the levels were further altered by SJG (Supplementary Fig. S1C). Then, gel electrophoresis for mass spectrometry (MS) was performed using 600 μg of a mixture of all liver protein samples by using the same above mentioned procedure. After fixation and staining of the gels, selected proteins were obtained from the gels by using an automated spot picker (GE Healthcare), and the gels were washed with 50 mM ammonium bicarbonate and acetonitrile. Then, the gels were incubated with 12.5 ng/μl trypsin (Promega) at 30 °C for 15 h. The peptides were eluted from the gels with 50 mM ammonium bicarbonate, followed by 10% (v/v) formic acid and acetonitrile. The solvent was changed to 0.1% (v/v) formic acid. The peptides were concentrated and desalted using the C18 PepMap column (300-μm id/5 mm in length; LC Packings, San Francisco, CA), followed by MS/MS sequencing with an automated MS-to-MS/MS switching protocol. Spectrum processing and peak annotations were performed, and processed spectra were searched using the Mascot database (Matrix Science, London, UK).

### 2.6. Statistical analysis

The microarray and proteome data are presented as means of log-normalized expression ± SEM. Differences between the groups were analyzed using Kruskal-Wallis analysis with Scheffe's multiple-comparison post-tests for non-normally distributed values or one-way analysis of variance (ANOVA) with Bonferroni's multiple-comparison post-tests for normally distributed values to evaluate the significance of the differences. A value of $P < 0.05$ was considered statistically significant.

## 3. Results

### 3.1. Effects of SJG on plasma and liver parameters

HCD rats were untreated or treated with SJG for 12 (short-term), 30 (medium-term), or 52 (long-term) weeks. Total cholesterol levels were highest in rats from the long-term experiment and lowest in rats from the short-term experiment, suggesting that cholesterol levels become elevated with age. HCD-induced hypercholesterolemia was improved by treatment with SJG in rats from the medium-term experiment (Table 1). Similar tendencies were observed for the short- and long-term experiments. Thus, we examined the genome-wide expression profiles in all three experiments. Treatment with SJG did not affect HCD-induced elevation of LDL cholesterol and free cholesterol levels. We also found that treatment with SJG did not affect glucose, TG, FFA, phospholipids, or total lipid levels. In addition, HCD induced fatty liver, and treatment with SJG did not ameliorate this effect (Table 1).

### 3.2. Effects of SJG on the genome-wide expression profile

We performed microarray analyses to identify gene sets whose expression was altered in the presence of of HCD, and then further affected by the presence of SJG. We identified only one gene in the long-term experiment by using Limma software (Supplementary Fig. S1A), and 44, 29, and 163 genes in the short-, medium-, and long-term experiments, respectively (Supplementary Fig. S1B), by using the Subio software. We did not identify any genes whose expression patterns were the identical or somewhat similar in the three experiments (Fig. 2).

### 3.3. Effects of SJG on the protein expression profile and comparison with the microarray results

We examined the protein expression profiles in the livers of rats from the long-term experiment because this group had the most affected genes in the standard microarray analysis (163 genes; Supplementary Fig. S1B). Then, we selected the proteins whose levels were altered by the intake of HCD, and were further affected by treatment with SJG, using a similar procedure for standard microarray analysis (Supplementary Fig. S1C), and identified paraoxonase 3 (PON3) and protein disulfide isomerase (PDIA1) as candidates (Fig. 3A and B, respectively). We found that expression of *PON3* and *PDIA1* transcripts (Fig. 3C and D, respectively) were comparable (Fig. 3A versus 3C; 3B versus 3D) to the protein expression. While PON3 protein levels and transcript levels of *PON3* and *PDIA1* in each experimental group showed normally distributed values, PDIA1 protein levels had non-normal distribution. The mean value and SEM of processed PON3 protein levels were 0.118 and 0.025 (MF+Water), −0.024 and 0.037 (MF+SJG), −0.106 and 0.067 (HCD+Water), and −0.019 and 0.032 (HCD+SJG), respectively. The mean value and SEM of processed *PON3* transcript levels were 0.014 and 0.026 (MF+Water), 0.028 and 0.026 (MF+SJG), −0.056 and 0.036 (HCD+Water), and 0.014 and 0.028 (HCD+SJG), respectively. The mean value and SEM of processed *PDIA1* transcript levels were 0.029 and 0.033 (MF+Water), 0.041 and 0.019 (MF+SJG), −0.056 and 0.021 (HCD+Water), and −0.015 and 0.015 (HCD+SJG), respectively. The median value, 25th percentile, and 75th percentile of processed PDIA1 protein levels were 0.084, 0.059, and 0.112 (MF+Water), −0.004, −0.029, and 0.025 (MF+SJG), −0.075, −0.146, and −0.056 (HCD+Water), and 0.045, 0.033, and 0.053 (HCD+SJG), respectively. We analyzed the PON3 protein levels and transcript levels of both genes using one-way ANOVA with Bonferroni's post-hoc test, and analyzed PDIA1 protein levels using Kruskal-Wallis test with Scheffe's post-hoc test. We found that protein levels of PON3 and PDIA1 were altered significantly between the groups while we did not observe any significant differences in transcript levels of

**Table 1**
Several parameters in plasma and liver.

| Sample | Parameters | Period | Experimental conditions | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | MF+water | | MF+SJG | | HCD+water | | HCD+SJG | |
| Blood | Glucose (mg/dl) | 12week | 110.4 ± 4.8 | a | 107.4 ± 3.4 | a | 105.0 ± 4.4 | a | 111.8 ± 7.4 | a |
| | | 30week | 122.8 ± 9.8 | a | 125.6 ± 6.8 | a | 115.0 ± 5.2 | a | 126.4 ± 4.2 | a |
| | | 52week | 111.6 ± 4.2 | a | 110.8 ± 7.2 | a | 111.0 ± 6.4 | a | 101.8 ± 1.2 | a |
| | Total cholesterol (mg/dl) | 12week | 66.0 ± 0.8 | a | 63.8 ± 3.0 | a | 93.8 ± 5.1 | b | 91.2 ± 5.2 | b |
| | | 30week | 73.6 ± 3.8 | a | 79.8 ± 6.3 | a | 112.6 ± 4.8 | b | 91.0 ± 4.3 | a |
| | | 52week | 108.2 ± 5.4 | a | 87.6 ± 8.6 | a | 124.6 ± 19.5 | a | 113.2 ± 3.9 | a |
| | HDL cholesterol (mg/dl) | 12week | 17.6 ± 0.4 | a | 17.8 ± 0.4 | a | 22.0 ± 0.5 | b | 19.2 ± 0.6 | a |
| | | 30week | 17.6 ± 0.7 | a | 18.0 ± 0.7 | a | 22.4 ± 1.1 | b | 18.0 ± 0.8 | a |
| | | 52week | 22.6 ± 0.8 | a | 20.8 ± 1.5 | a | 22.8 ± 3.5 | a | 23.0 ± 1.2 | a |
| | LDL cholesterol (mg/dl) | 12week | 5.8 ± 0.2 | a | 6.0 ± 0.4 | a | 18.4 ± 1.7 | b | 21.4 ± 2.0 | b |
| | | 30week | 6.0 ± 0.3 | a | 6.8 ± 0.7 | a | 20.8 ± 1.4 | b | 17.6 ± 1.1 | b |
| | | 52week | 12.8 ± 1.0 | a | 10.0 ± 0.9 | a | 29.2 ± 5.8 | b | 25.8 ± 1 | b |
| | Cholesteryl ester (mg/dl) | 12week | 56.2 ± 0.6 | a | 54.4 ± 2.3 | a | 79.8 ± 4.3 | b | 78.4 ± 4.5 | b |
| | | 30week | 54.6 ± 2.9 | a | 59.4 ± 4.5 | ac | 89.4 ± 3.8 | b | 71.2 ± 3.5 | c |
| | | 52week | 89.0 ± 4.7 | a | 74.2 ± 6.8 | a | 104.0 ± 16.5 | a | 95.8 ± 3.4 | a |
| | Free cholesterol (mg/dl) | 12week | 9.8 ± 0.4 | a | 9.4 ± 0.7 | a | 14.0 ± 0.8 | b | 12.8 ± 0.7 | b |
| | | 30week | 19.0 ± 0.9 | a | 20.4 ± 1.9 | a | 23.2 ± 1.0 | a | 19.8 ± 0.9 | a |
| | | 52week | 19.2 ± 1.2 | a | 13.4 ± 1.8 | a | 20.6 ± 3.0 | a | 17.4 ± 0.5 | a |
| | TG (mg/dl) | 12week | 128.6 ± 18.4 | a | 109.8 ± 15.4 | a | 86.2 ± 6.8 | a | 88.4 ± 7.3 | a |
| | | 30week | 161.4 ± 29.4 | a | 148.4 ± 18.6 | a | 121.8 ± 9.2 | a | 103.8 ± 13.2 | a |
| | | 52week | 290.2 ± 38.8 | a | 218.2 ± 20.5 | ac | 156.8 ± 22.2 | bc | 176.2 ± 9.2 | bc |
| | Free fatty acid (μEq/l) | 12week | 839.6 ± 59.7 | a | 838.6 ± 69.4 | a | 738.4 ± 55.7 | a | 809.0 ± 79.1 | a |
| | | 30week | 787.2 ± 42.4 | a | 845.0 ± 74.7 | a | 911.8 ± 65.4 | a | 736.0 ± 45.8 | a |
| | | 52week | 964.6 ± 54.7 | a | 981.0 ± 46.7 | a | 869.2 ± 89.0 | a | 928.6 ± 52.3 | a |
| | Phospholipids (mg/dl) | 12week | 113.2 ± 2.8 | a | 109.4 ± 4.3 | a | 113.2 ± 3.5 | a | 106.0 ± 4.0 | a |
| | | 30week | 119.8 ± 4.8 | a | 125.2 ± 8.1 | a | 125.6 ± 5.0 | a | 111.6 ± 5.2 | a |
| | | 52week | 166.4 ± 7.6 | a | 140.6 ± 10.8 | a | 131.4 ± 14.9 | a | 125.8 ± 5.1 | a |
| | Total lipids (mg/dl) | 12week | 308.4 ± 17.2 | a | 287.0 ± 18.8 | a | 337.8 ± 15.7 | a | 344.4 ± 22.5 | a |
| | | 30week | 391.0 ± 29.5 | a | 392.2 ± 32.9 | a | 451.6 ± 14.5 | a | 383.0 ± 17.7 | a |
| | | 52week | 563.4 ± 37.1 | a | 484.0 ± 35.4 | a | 501.4 ± 58.9 | a | 499.6 ± 15.6 | a |
| Liver | Total cholesterol (mg/100 g) | 12week | 506.8 ± 9.8 | a | 488.0 ± 8.8 | a | 3582.5 ± 271.0 | b | 4119.3 ± 298.5 | b |
| | | 30week | 504.3 ± 15.0 | a | 512.8 ± 6.9 | a | 4302.0 ± 190.9 | b | 3988.7 ± 203.1 | b |
| | | 52week | 760.2 ± 20.6 | a | 739.6 ± 20.1 | a | 5112.5 ± 373.1 | b | 4303.4 ± 154.5 | b |

Each value is the mean ± SEM of four individuals. Different letters (a, b, and c) denote a statistically significant difference in means ($P < 0.05$; one-way ANOVA with Bonferroni's multiple-comparison post-tests for multiple comparisons).

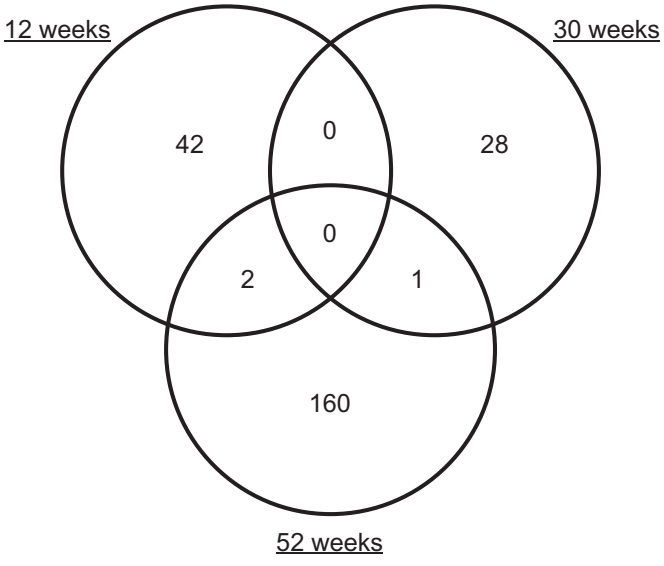


**Fig. 2.** Results of genome-wide expression analysis of the liver of diet-induced hypercholesterolemic rats. Venn diagrams showing the number of genes altered by cholesterol or a combination of cholesterol and SJG in the short- (12 weeks), medium- (30 weeks), and long-term (52 weeks) experiments. The total number of genes whose expressions were altered by HCD intake and further modified by SJG was 44, 29, and 163 in the short-, medium-, and long-term experiments, respectively. No genes showed similar expression patterns in all three experiments.

either gene. SJG was potent enough to recover HCD-induced reduction of PON3 and PDIA1 levels.

We searched for data about *PDIA1* in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. This gene is associated with "protein processing in endoplasmic reticulum" (accession number, rno04141). We collected expression data for 30,451 transcripts by using microarray. Of the 30,451 genes, 184 were associated with the rno04141 pathway, and 3998 had expression patterns that were similar to that of *PDIA1* (Pearson's correlation, $r > 0.85$, Table 2). Of the 3998 genes, 37 genes (Supplementary Table S1) were associated with the rno04141 pathway (Table 2). By analyzing these data with the Fisher's exact test, we found that the genes with expression patterns similar to that of *PDIA1* were significantly enriched in the rno04141 pathway (Table 2, $P < 0.01$). These results suggest that intake of HCD reduces the expression of genes in this pathway and that SJG may play a role in the recovery of that expression.

PON3 is a member of the arylesterase/paraoxonase family, and it acts as an anti-atherosclerotic agent by binding to HDL [18]. We could not find atherosclerosis-related pathways in the KEGG database. Thus, we examined the expression profiles of genes related to the formation of atherosclerotic lesions, such as *PON1*, *PON2*, *PON3*, matrix metalloproteinase 9 (*MMP9*), scavenger receptor AI (*SR-AI*), *CD36*, lectin-type oxidized LDL receptor 1 (*LOX1*), and endothelial nitric oxide synthase (*eNOS*), and found that HCD-induced changes in the expression of these genes tended to improve with SJG intake (Fig. 4). However, the expression of several other genes with possible roles in the formation

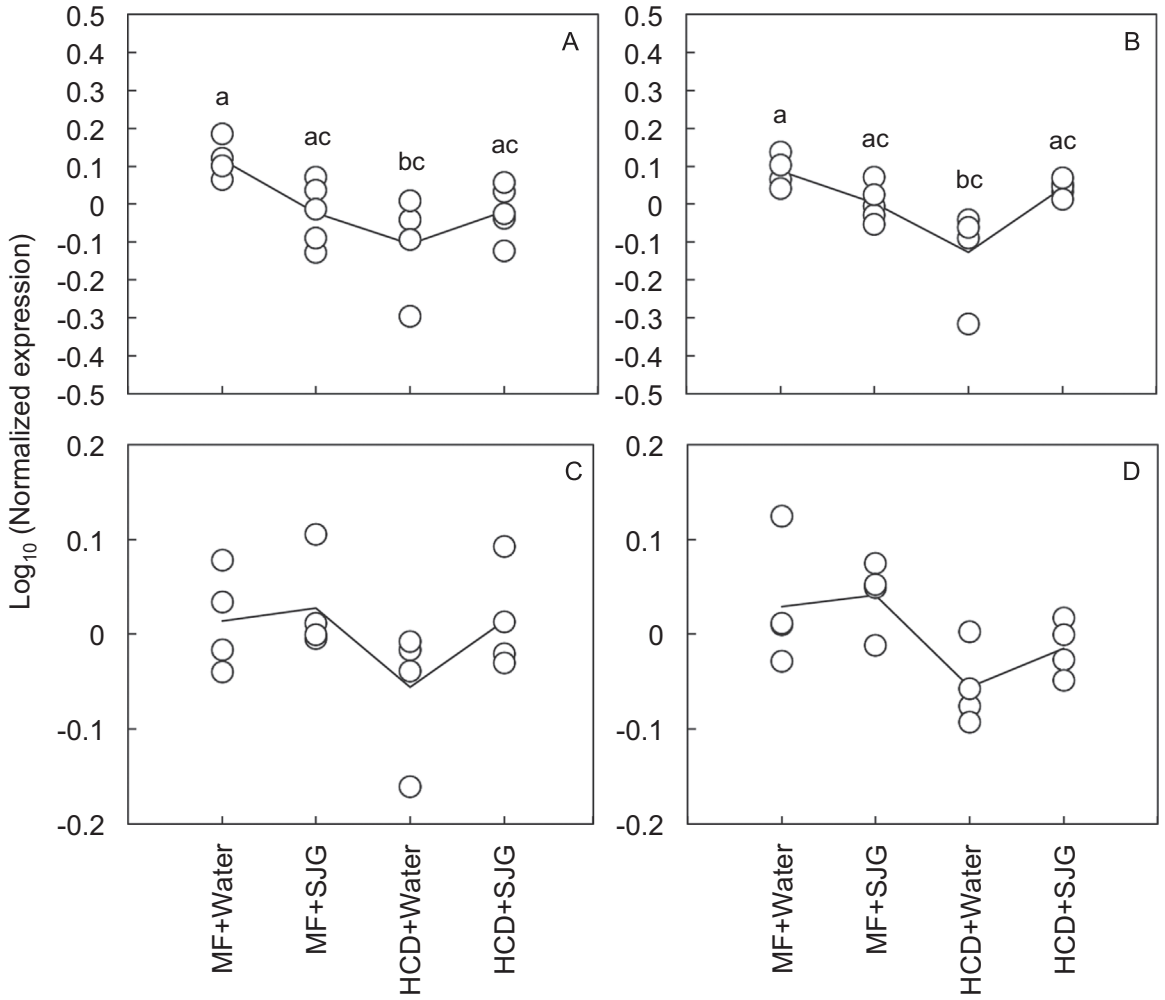


**Fig. 3.** Results of proteome and transcriptome analyses of the liver of diet-induced hypercholesterolemic rats. Two-dimensional gel electrophoresis of rat liver proteins was performed using immobilized pH 3–10 gradient strips, followed by SDS-PAGE (12% T, 7.5% C). Panels A and B represent quantified data for proteome analysis of PON3 and PDIA1, respectively. Panels C and D represent quantified data for transcriptome analysis of PON3 and PDIA1, respectively. The vertical axis represents the standardized quantity of the target protein or transcript for each sample (n=5 for proteome analysis and n=4 for transcriptome analysis). The open circle represents the standardized quantity of each sample. The line represents the average value for each treatment group. Different letters denote a statistically significant difference in means ($P<0.05$; Kruskal-Wallis analysis with Scheffe's multiple-comparison post-tests or one-way ANOVA with Bonferroni's multiple comparison post-test for multiple comparison).

of atherosclerotic lesions, such as intercellular adhesion molecule 1 (*ICAM1*), vascular cell adhesion molecule 1 (*VCAM1*), C-C chemokine receptor type 2, macrophage colony-stimulating factor (*MCSF*), monocyte chemotactic protein-1, interleukin-6, and tumor necrosis factor alpha (*TNF*-α), were not affected (data not shown).

## 4. Discussion

In the present study, we aimed to establish a combined method for analyzing data from transcriptome experiments coupled with proteome analyses, which enabled us to identify genes of interest that could not be selected using standard methods. Crude herbal drugs are usually prepared from several plants and are thought to have mild effects when compared to the effects of pharmaceutical agents. Thus, changes caused by treating organisms with herbal crude drugs may not be large enough, making it difficult to obtain useful information from genome-wide experiments using standard methods. First, we examined several serum parameters and found that total cholesterol and HDL cholesterol levels were elevated by HCD intake and that the HCD-induced increases of total cholesterol or HDL cholesterol were recovered by SJG in the

**Table 2**
Pathway enrichment analysis using Fisher's exact test.

| | Number of genes associated with the KEGG pathway rno04141 | Number of genes not associated with the KEGG pathway rno04141 | Total |
|---|---|---|---|
| Number of genes whose expression was similar to PDIA1 | 37 | 3961 | 3998 |
| Number of genes whose expression was not similar to PDIA1 | 147 | 26,306 | 26,453 |
| Total | 184 | 30,267 | 30,451 |

Genes whose expression was similar to that of PDIA1 were selected according to Peason's correlation ( $>0.85$), and were enriched significantly in KEGG pathway rno04141 (protein processing in endoplasmic reticulum). $P<0.01$ (Fisher's exact test, two-tailed).

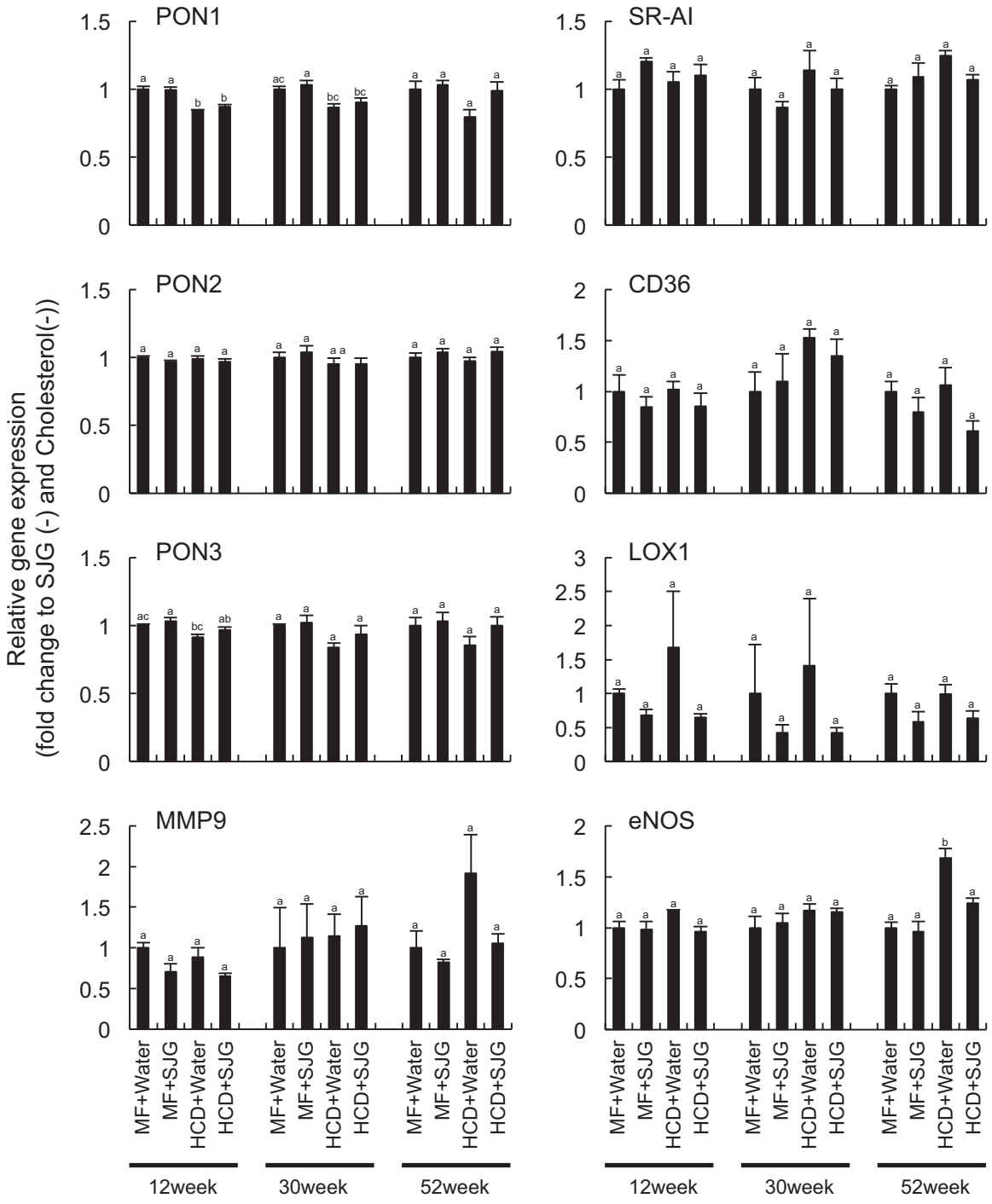


**Fig. 4.** Expression profiles of the genes related to the formation of atherosclerotic lesions. Data were extracted from the microarray results. The panels represent the expression profiles of PON1 (A), PON2 (B), PON3 (C), MMP9 (D), SR-AI (E), CD36 (F), LOX1 (G), and eNOS (H). The vertical axis represents the ratio of the quantity of the target gene transcript for each sample to that of group 1 samples as a magnitude of induction (fold change) after global normalization (n=4). Each value is mean ± SEM. Different letters denote a statistically significant difference in means ($P < 0.05$; one-way ANOVA with Bonferroni's multiple-comparison post-tests for multiple comparisons).

medium-tern experiment or short- and medium-term experiments, respectively. Although the effect of SJG was not consistent among the short-, medium-, and long-term experiments, it is possible that SJG has beneficial effects on cholesterol homeostasis in liver.

Next, we attempted to extract gene sets from genome-wide expression profiles using standard methods. However, we could not find gene(s) with expression commonly altered in short-, medium-, and long-term experiments (Fig. 2). Additionally, *PON3* and *PDIA1*, which were not selected by standard methods, were selected from proteome analysis. These results suggested that, at least in this study, it is difficult to obtain useful candidates using standard methods. Therefore, we performed proteome analysis of samples from the long-term experiment and found that PON3 and PDIA1 levels were reduced by HCD intake and that SJG recovered the levels (Fig. 3A and B). Gene expression patterns of *PON3* and *PDIA1* in the long-term experiment were comparable to the protein expression patterns (Fig. 3C and D). Based on these results, it is evident that we were unable to identify *PON3* and *PDIA1* as genes of interest using standard methods. However, we could identify PON3 and PDIA1 as candidate proteins from

the proteome analyses and found a correlation between the transcript and protein levels of both these genes. There might be several cases in which expression levels are different between the protein and the transcript because the protein functions are altered due to post-translational modifications like acetylation, methylation, or phosphorylation. Although, in such cases, it will be more difficult to identify the genes of interest from transcriptomic data, the combined analytical method is considered a powerful tool because it provides a large amount of data.

We demonstrated that genes with expression patterns similar to that of *PDIA1* were significantly enriched in the "protein processing in endoplasmic reticulum" pathway, suggesting that protein processing was decreased by HCD intake and that SJG had the ability to recover the HCD-induced decrease, even if the magnitude of the change was small (Table 2).

*PON1* and *PON3* had similar expression patterns (Fig. 4). PON1 and PON3 are synthesized mainly in the liver and bind to HDL [18]. PON1-HDL and PON3-HDL have antioxidant activity and prevent the oxidation of LDL during oxidative stress [19,20]. Oxidized LDL plays a role in the occurrence or progression of atherosclerosis by acting as a signal molecule. Oxidative stress increases the migration of monocytes and upregulates the expression of adhesion molecules, including ICAM1 and VCAM1 [21]. After binding to adhesion molecules, monocytes are stimulated by MCSF to move into the subendothelial region [22]. Then, the monocytes are further stimulated by oxidized LDL in the subendothelial region, and they differentiate into macrophages [22]. Scavenger receptors, such as SR-AI, CD36, and LOX1, play an important role in these steps [22]. In addition, inflammation, migration of smooth muscle cells, and interactions with the extracellular matrix increase cytokine levels, which causes the macrophages to form foam cells [23–28]. MMP9 induces the rupture of atherosclerotic lesions and causes ischemic heart disease [22,29]. Thus, the effects of SJG on HCD-induced reduction of *PON1* and *PON3* may lead to the improvement of atherosclerosis. The expression of genes related to the formation of atherosclerosis, including scavenger receptors, adhesion molecules, and MMPs, was improved after treatment with SJG.

In conclusion, we established a combined method for the analysis of transcriptome experiments by coupling with them with proteome experiments, and we identified genes of interest that were not identified using standard methods. These methods will be useful in identifying gene sets that do not show large changes in expression.

## Acknowledgements

This work was supported in part by Grant-in-Aid for Young Scientists (B) (25870296) from Japan Society for Promotion of Science, Japan to AI. We are grateful to Mr. Akio Tanabe, who is a President of Subio Inc. and is professional of statistics and analysis of genome-wide expression profiles, for a consulting of data analysis in this study.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.bbrep.2016.05.015.

## References

[1] N. Kaminski, N. Friedman, Practical approaches to analyzing results of microarray experiments, Am. J. Respir. Cell Mol. Biol. 27 (2002) 125–132.
[2] X. Cui, G.A. Churchill, Statistical tests for differential expression in cDNA microarray experiments, Genome Biol. 4 (2003) 210.
[3] R. Nadon, J. Shoemaker, Statistical issues with microarrays: processing and analysis, Trends Genet. 18 (2002) 265–271.
[4] G.C. Tseng, M.K. Oh, L. Rohlin, J.C. Liao, W.H. Wong, Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects, Nucleic Acids Res. 29 (2001) 2549–2557.
[5] C. Gardmo, A. Tamburro, S. Modica, A. Moschetta, Proteomics for the discovery of nuclear bile acid receptor FXR targets, Biochim. Biophys. Acta 2011 (1812) 836–841.
[6] T. Kondo, Tissue proteomics for cancer biomarker development: laser microdissection and 2D-DIGE, BMB Rep. 41 (2008) 626–634.
[7] J. Armengaud, J. Trapp, O. Pible, O. Geffard, A. Chaumot, E.M. Hartmann, Non-model organisms, a species endangered by proteogenomics, J. Proteom. 105 (2014) 5–18.
[8] A. Ishihara, R. Takahashi, T. Numajiri, S. Kaneko, Y. Ishizawa, S. Koya, K. Yamauchi, Effects of Japanese herbal crude drug, combined extract of Sasa albo-marginata leaves, Japanese red pine leaves and ginseng roots on gene expression in hypercholesterolemic rat liver, J. Tradit. Med. 30 (2013) 236–245.
[9] W.M. Loke, J.M. Proudfoot, J.M. Hodgson, A.J. McKinley, N. Hime, M. Magat, R. Stocker, K.D. Croft, Specific dietary polyphenols attenuate atherosclerosis in apolipoprotein E-knockout mice by alleviating inflammation and endothelial dysfunction, Arterioscler. Thromb. Vasc. Biol. 30 (2010) 749–757.
[10] C. Peluzio Mdo, T.F. Teixeira, V.P. Oliveira, C.M. Sabarense, C.M. Dias, M.V. Abranches, I.R. Maldonado, Grape extract and alpha-Tocopherol effect in cardiovascular disease model of Apo E -/- Mice, Acta Cir. Bras. 26 (2011) 253–260.
[11] E.V. Shabrova, O. Tarnopolsky, A.P. Singh, J. Plutzky, N. Vorsa, L. Quadro, Insights into the molecular mechanisms of the anti-atherogenic actions of flavonoids in normal and obese mice, PLoS One 6 (2011) e24634.
[12] J.H. Suh, A. Virsolvy, A. Goux, C. Cassan, S. Richard, J.P. Cristol, P.L. Teissedre, J. M. Rouanet, Polyphenols prevent lipid abnormalities and arterial dysfunction in hamsters on a high-fat diet: a comparative study of red grape and white persimmon wines, Food Funct. 2 (2011) 555–561.
[13] L. Verschuren, P.Y. Wielinga, W. van Duyvenvoorde, S. Tijani, K. Toet, B. van Ommen, T. Kooistra, R. Kleemann, A dietary mixture containing fish oil, resveratrol, lycopene, catechins, and vitamins E and C reduces atherosclerosis in transgenic mice, J. Nutr. 141 (2011) 863–869.
[14] H. Kuribara, H. Iwata, H. Tomioka, R. Takahashi, K. Goto, N. Murohashi, S. Koya, The anxiolytic effect of Sho-ju-sen, a Japanese herbal medicine, assessed by an elevated plus-maze test in mice, Phytother. Res. 15 (2001) 142–147.
[15] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (2015) e47.
[16] H. Yamanaka, Y. Yakabe, K. Saito, M. Sekijima, T. Shirai, Quantitative proteomic analysis of rat liver for carcinogenicity prediction in a 28-day repeated dose study, Proteomics 7 (2007) 781–795.
[17] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, Anal. Biochem. 72 (1976) 248–254.
[18] M. Aviram, J. Vaya, Paraoxonase 1 activities, regulation, and interactions with atherosclerotic lesion, Curr. Opin. Lipidol. 24 (2013) 339–344.
[19] N.Y. Gbandjaba, N. Ghalim, M. Hassar, H. Berrougui, H. Labrazi, H. Taki, R. Saile, A. Khalil, Paraoxonase activity in healthy, diabetic, and hemodialysis patients, Clin. Biochem. 45 (2012) 470–474.
[20] L.P. Precourt, D. Amre, M.C. Denis, J.C. Lavoie, E. Delvin, E. Seidman, E. Levy, The three-gene paraoxonase family: physiologic roles, actions and regulation, Atherosclerosis 214 (2011) 20–36.
[21] H. Chen, D. Li, T. Saldeen, J.L. Mehta, Transforming growth factor-beta(1) modulates oxidatively modified LDL-induced expression of adhesion molecules: role of LOX-1, Circ. Res. 89 (2001) 1155–1160.
[22] T. Goyal, S. Mitra, M. Khaidakov, X. Wang, S. Singla, Z. Ding, S. Liu, J.L. Mehta, Current concepts of the role of oxidized LDL receptors in atherosclerosis, Curr. Atheroscler. Rep. (2012).
[23] Y. Sun, X. Chen, Ox-LDL-induced LOX-1 expression in vascular smooth muscle cells: role of reactive oxygen species, Fundam. Clin. Pharmacol. 25 (2011) 572–579.
[24] C. Hu, A. Dandapat, L. Sun, J.A. Khan, Y. Liu, P.L. Hermonat, J.L. Mehta, Regulation of TGFbeta1-mediated collagen formation by LOX-1: studies based on forced over-expression of TGFbeta1 in wild-type and lox-1 knock-out mouse cardiac fibroblasts, J. Biol. Chem. 283 (2008) 10226–10231.
[25] C. Hu, A. Dandapat, L. Sun, J. Chen, M.R. Marwali, F. Romeo, T. Sawamura, J.L. Mehta, LOX-1 deletion decreases collagen accumulation in atherosclerotic plaque in low-density lipoprotein receptor knockout mice fed a high-cholesterol diet, Cardiovasc. Res. 79 (2008) 287–293.
[26] J.L. Mehta, N. Sanada, C.P. Hu, J. Chen, A. Dandapat, F. Sugawara, H. Satoh, K. Inoue, Y. Kawase, K. Jishage, H. Suzuki, M. Takeya, L. Schnackenberg, R. Beger, P. L. Hermonat, M. Thomas, T. Sawamura, Deletion of LOX-1 reduces atherogenesis in LDLR knockout mice fed high cholesterol diet, Circ. Res. 100 (2007) 1634–1642.
[27] D. Li, J.L. Mehta, Upregulation of endothelial receptor for oxidized LDL (LOX-1) by oxidized LDL and implications in apoptosis of human coronary artery endothelial cells: evidence from use of antisense LOX-1 mRNA and chemical inhibitors, Arterioscler. Thromb. Vasc. Biol. 20 (2000) 1116–1122.
[28] D. Li, J.L. Mehta, Antisense to LOX-1 inhibits oxidized LDL-mediated upregulation of monocyte chemoattractant protein-1 and monocyte adhesion to human coronary artery endothelial cells, Circulation 101 (2000) 2889–2895.
[29] D. Li, L. Liu, H. Chen, T. Sawamura, S. Ranganathan, J.L. Mehta, LOX-1 mediates oxidized low-density lipoprotein-induced expression of matrix metalloproteinases in human coronary artery endothelial cells, Circulation 107 (2003) 612–617.