Research Article

Cytotoxic Effect of Recombinant *Mycobacterium tuberculosis* CFP-10/ESAT-6 Protein on the Crucial Pathways of WI-38 Cells

Kun-Nan Tsai,[1] Err-Cheng Chan,[2] Tsung-Yeh Tsai,[1] Kuei-Tien Chen,[2] Chun-Yu Chen,[2] Kenneth Hung,[1] and Chung-Ming Chen[1]

[1] *Institute of Biomedical Engineering, National Taiwan University, 1, Section 1, Jen-Ai Rd, Taipei 100, Taiwan*
[2] *Department of Medical Biotechnology and Laboratory Science, Chang Gung University, 259 Wen-Hua 1st Road, Kweishan, Taoyuan 333, Taiwan*

Correspondence should be addressed to Chung-Ming Chen, chung@ntu.edu.tw

To unravel the cytotoxic effect of the recombinant CFP-10/ESAT-6 protein (rCFES) on WI-38 cells, an integrative analysis approach, combining time-course microarray data and annotated pathway databases, was proposed with the emphasis on identifying the potentially crucial pathways. The potentially crucial pathways were selected based on a composite criterion characterizing the average significance and topological properties of important genes. The analysis results suggested that the regulatory effect of rCFES was at least involved in cell proliferation, cell motility, cell survival, and metabolisms of WI-38 cells. The survivability of WI-38 cells, in particular, was significantly decreased to 62% with 12.5 μM rCFES. Furthermore, the focal adhesion pathway was identified as the potentially most-crucial pathway and 58 of 65 important genes in this pathway were downregulated by rCFES treatment. Using qRT-PCR, we have confirmed the changes in the expression levels of LAMA4, PIK3R3, BIRC3, and NFKBIA, suggesting that these proteins may play an essential role in the cytotoxic process in the rCFES-treated WI-38 cells.

1. Introduction

Mycobacterium tuberculosis is a global infectious disease that has affected one-third of the world population, killing 2-3 million people and causing 7-8 million new infections annually [1]. *M. tuberculosis* (MTB) is a major cause of human tuberculosis. During the early stages of human tuberculosis, MTB induces an immune response [2] and subsequently leads to the development of lung granulomas consisting of macrophages, T cells, B cells, and fibroblasts [3]. Recent researches reveal that fibroblasts are not only essential in secreting chemokine for modulating inflammatory response to MTB infection and influencing the survival of MTB within macrophages [4] but also involved in the regulation of granuloma formation during MTB infection [4, 5]. Despite the potentially vital role of fibroblasts in MTB infection, the detailed MTB-regulated mechanism in fibroblasts, especially its relationship to MTB secreted protein, remains unknown.

Two secreted proteins CFP-10 and ESAT-6, produced by the region of difference 1 (RD1) in MTB, have been identified to play important roles in the pathogenesis of tuberculosis [6–8] in primary pulmonary infection. These two proteins have also been shown to be virulence factors with cytotoxic effects on macrophages, lung epithelial cells, and dendritic cells [8–10]. Individually, the cytotoxic effect of ESAT-6 protein has been found to evoke apoptosis of macrophages, dendritic cells, and fibroblasts [11]. Recent report reveals that CFP-10 and ESAT-6 choose a stable structure, forming a 1 : 1 heterodimeric complex [12]—the CFP-10/ESAT-6 protein (CFES). It has been shown that CFES elicits immune response in the host organism [13, 14]. However, the role and function of CFES in fibroblasts is not clear. Therefore, as our first attempt to unravel the effect of CFES on fibroblasts, an integrative analysis approach combining biological resources and bioinformatics was developed in this study.

Microarray is a biological resource, which has been often used to analyze gene expression profiles in biological

experiments. Most analysis tools for microarray data, for example, SAM [15], LPE [16], Bayesian [17], and so forth, have been designed mainly for identification of important genes. Other tools like GenMAPP [18], PharmGKB [19], and KEGG [20] only show the positions of the genes on a known pathway. Although some softwares such as ArrayXpath [21] integrate the pathway resources and provide analysis and visualization tools for deciphering the important genes and pathway structures, no notion of chaining or aggregating regulatory effect in a biological process has been taken into account in identifying the crucial pathways. To characterize the regulation mechanism of recombinant CFP-10/ESAT-6 protein (rCFES) on WI-38 cells taking account of the chaining or aggregating regulatory effect, an integrative analysis approach was proposed in this study. By combining time-course microarray data and annotated pathway databases, a new composite score quantifying the average significance and topological properties of important genes in a pathway was proposed to identify the potentially crucial pathways. Biologically, a crucial pathway in this study is a pathway that is substantially influenced by the rCFES-treatment, the consequence of which is highly related to the observed response in the rCFES-treated WI-38 cells, for example, the increased cell death rate. Nevertheless, since the crucial pathways suggested by any computational analysis approach like the proposed one require further experimental verification, they are thus considered as "potentially crucial pathways". A pathway with a better composite score was considered to be potentially more crucial in an rCFES-treated WI-38 cell in the sense that (i) the important genes in this pathway were more significantly expressed, (ii) this pathway contained a higher percentage of important genes, and (iii) the important genes in this pathway interacted more closely with one another. Based on the composite scores, the pathways with the best composite scores were suggested as the potentially crucial pathways in an rCFES-treated WI-38 cell, which may serve as the basis for further experimental studies on unraveling the cytotoxic effect of rCFES on WI-38 cells.

2. Materials and Methods

2.1. Cell Cultures. WI-38 cells were cultured in Modified Eagle Medium (Gibco) containing 10% fetal calf serum (FCS), 2 mM L-glutamine, 1 mM sodium pyruvate, 100 μg/mL streptomycin, 0.025 μg/mL amphotericin B, and 100 U/mL penicillin at 37°C with 5% CO2.

2.2. Expression and Purification of rCFES. The bacterial vector pET29b-CFES was created by cloning CFP-10 and ESAT-6 from the H37Rv strain of MTB and was transformed into *Escherichia coli* BL21 (DE3). IPTG was used to express abundant amounts of recombinant CFP-10/ESAT-6 protein (rCFES). Next, we purified recombinant rCFES using affinity chromatography with nickel ion characteristic and dialysis. The Bio-Rad DC Protein Assay Kit was used to measure rCFES concentrations. The purity of rCFES was controlled at a level higher than 95% as assessed by densitometry of 12% SDS-PAGE gels. In addition, MALTI-TOF mass spectroscopy was used to confirm that rCFES was composed of CFP-10 and ESAT-6. More detailed confirmation analysis may be found in Supplemental Data I (http://homepage.ntu.edu.tw/~d91548013/CFES_supp_data1.pdf).

2.3. Cell Survival Assay. WI-38 cells were seeded at 1.3×10^4 cells per well in 96-well plates containing Modified Eagle Medium with 10% FCS. The fibroblasts were arrested at the G0/G1 phase after treatment with serum-free medium for 24 hours, further treated with rCFES at different concentrations, including 0, 0.625, 1.25, 3.125, 6.25, 12.5, and 25 μM, and incubated for 48 hours. The cell viability was measured with Cell Proliferation Assay Kit (Promega). The absorbance was recorded at 490 nm using a 96-well-plated reader. The assay was performed in triplicate on control and rCFES-treated WI-38 cells at different concentrations.

2.4. Microarray Analysis. WI-38 cells were arrested at the G0/G1 phase after treatment with serum-free medium for 24 hours, treated with rCFES (12.5 μM), and incubated for 0, 3, 8, 16, 24, 32, 40, and 48 hours. Total RNA was extracted with Trizol reagent, according to the manufacturer's protocol (Invitrogen) and with RNeasy Mini Kit (Qiagen). RNA was also extracted from nontreated WI-38 cells as control at the same times. In this study, experiments on control and rCFES-treated cells were performed in triplicate at each time point. Purified RNA was quantified at OD260 with an ND-1000 spectrophotometer (Nanodrop) and qualitated with Agilent Bioanalyzer 2100 (Agilent).

Total RNA was amplified with a Fluorescent Linear Amplification Kit (Agilent) and labelled with Cy3-CTP or Cy5-CTP (CyDye, PerkinElmer). Next, cRNA was hybridized onto human whole genome oligo microarrays (Agilent), according to the manufacturer's protocols. A total of 43 931 probe sets on the arrays were analyzed. After washing and drying by blowing with a nitrogen gun, microarrays were scanned with an Agilent microarray scanner (Agilent) at 535 nm for Cy3 and 625 nm for Cy5. Feature Extraction Software (Agilent) was used to analyze the scanned images and to estimate differential gene expression by calculating statistical confidences. We selected 41 675 probe sets by rank-consistency-filtering using the LOWESS method.

2.5. Significance Analysis of Gene Expression. The following is screening conditions for selecting the matched genes at every time point: rBGSubSignal >32, gBGSubSignal >32, and absolute value log2 ratio >0. The unions of genes passing this filter were used as the input data. Significance analysis of microarray (SAM) [15] was then performed to identify important genes using the false discovery rate (FDR) controlling procedures.

2.6. Pathway Topology Analysis. Pathway topology analysis aimed to identify the potentially crucial pathways involved in the biological process for WI-38 cells treated with rCFES. The potentially crucial pathways were selected in two steps

based on a composite criterion characterizing multiple clustering properties of important genes. In the first step, called pathway enrichment analysis, a set of pathways were chosen as the candidates of the potentially crucial pathways, each of which consisted of a significant proportion of genes that were important genes in the pathway. In the second step, these candidates were ranked according to a scoring system combining several descriptors that quantify the average significance and topological properties of important genes.

More specifically, in the first step, important genes were mapped onto the human pathways constructed in the KEGG databases using LocusLink IDs. Supposed that there were T total genes with LocusLink IDs, out of which N genes were mapped to a pathway, say pathway P_j. Moreover, supposed that there were t important genes in total, out of which k important genes were involved in pathway P_j. To determine if pathway P_j contained a significant proportion of important genes, hyper geometric statistic test was used with the null hypothesis that k/t was not greater than N/T, given that N and t are both greater than 0. The P-value for pathway P_j was calculated as

$$p = 1 - \sum_{i=0}^{k} \frac{\binom{t}{i}\binom{T-t}{N-i}}{\binom{T}{N}}. \tag{1}$$

While the P-value represented the significance of pathway P_j having a significant proportion of genes that were important genes, it also stood for the Type I error rate, that is, the probability that k/t was mistakenly inferred to be significantly higher than N/T. Statistically, the more pathways were tested, the greater probability it might have to observe a false-positive result. With the hundreds of pathways to be tested, to control the overall Type I error rate of multiple testing, false discovery rate (FDR) [22] correction was employed in this study.

The second step was to further rank the importance of the candidates selected in the first step based on the average significance and topological properties of important genes. The average significance quantified the overall significance of the important genes identified in pathway P_j. Two descriptors were computed to represent the average significance. The first one, denoted by Q_s, was defined as the geometric mean of the q-values computed by SAM for the important genes in pathway P_j. The geometric mean instead of arithmetic mean was adopted because the q-values of the genes in a pathway may span several orders of magnitude. If arithmetic mean is taken, the genes with large q-values, even if they are minority in the pathway, may dominate the mean value, yielding a misleading result. The second one, denoted by Q_{ps}, was defined as the geometric mean of the cooccurrence significances of all pairs of important genes in pathway P_j. A pair of important genes was composed of two directly connected important genes in the KEGG pathway database and the cooccurrence significance of a pair of important genes was defined as the product of the q-values of these two genes. While Q_s measured the overall significance due to individual important genes, Q_{ps}

placed more emphasis on the significance characterizing the cooccurrence of a pair of important genes. It was assumed that a gene pair with two highly important genes would be more valuable than that with two genes of moderate significance.

Two classes of topological properties of important genes were quantified to represent the overall significance of pathway P_j in the underlying biological process. One was density and the other was clustering. Density referred to the average compactness of important genes in pathway P_j. Two density descriptors, named Q_d and Q_{pd}, were computed. Q_d was defined as the ratio of the number of important genes to the total number of genes in pathway P_j, whereas Q_{pd} was defined as the ratio of the number of pairs of important genes to the total number of gene pairs in the pathway. A gene pair consisted of two directly connected genes in the KEGG pathway database.

Clustering was performed based on the important genes mapped onto the KEGG database and the connection information among the mapped important genes provided by the KEGG database. Clustering revealed how closely the important genes were connected. A cluster was an aggregation of connected important genes. Two types of clustering, that is, directed clustering and undirected clustering, were considered, which were characterized by two pairs of topological descriptors. In a directed cluster, one may traverse all important genes starting from at least one gene, called root gene, following the upstream-downstream relation between every pair of connected important genes. On the other hand, in an undirected cluster, every link connecting two important genes defined in the KEGG databases was regarded as an undirected link. That is, one may traverse from any gene in an undirected cluster to all the other genes in the same cluster without considering the upstream-downstream relation. More detailed descriptions for the roles of the directed and undirected clusterings may be found in Supplemental Data I (http://homepage.ntu.edu.tw/~d91548013/CFES_supp_data1.pdf).

The first pair of topological descriptors, denoted as N_{cd} and Q_{cd}, stood for the number of important genes and the geometric mean of the q-values computed by SAM for the important genes, respectively, in the maximum directed cluster of pathway P_j. If more than one maximum directed cluster was identified in pathway P_j, the maximum directed cluster with the smallest Q_{cd} was chosen. The second pair of topological descriptors, denoted as N_{cu} and Q_{cu}, represented the number of important genes and the geometric mean of the q-values of the important genes, respectively, in the maximum undirected cluster of pathway P_j. Similarly, if more than one maximum undirected cluster was identified in pathway P_j, the maximum undirected cluster with the smallest Q_{cu} was chosen. For convenience, the clusters quantified by the first, and second pairs of topological descriptors were termed as MD-cluster, and MUD-cluster, respectively.

To determine the potentially crucial pathways, the candidate pathways selected in the first step were ranked based on a composite score, R_r, which was defined as the sum of the individual ranks of the descriptors for the average

significance and topological properties. More specifically, $R_r = R_a + R_d$, where

$$R_a = \text{rank}(S_a(Q_s)) + \text{rank}\left(S_a\left(Q_{ps}\right)\right)$$
$$+ \text{rank}(S_a(Q_{cd})) + \text{rank}(S_a(Q_{cu})),$$
$$R_d = \text{rank}(S_d(Q_d)) + \text{rank}\left(S_d\left(Q_{pd}\right)\right)$$
$$+ \text{rank}(S_d(N_{cd})) + \text{rank}(S_d(N_{cu})),$$

(2)

where $S_a(X)$ and $S_d(Y)$ are two sorting functions that sort X and Y in the ascending order and descending order, respectively, and "*rank*" denotes the rank in the sorted sequence. The smaller the composite score was, the more likely a pathway was considered to be a potentially crucial pathway.

2.7. QRT-PCR Analysis. To validate the microarray data, real-time qRT-PCR using SYBR Green I Kit (Roche) was performed to assess RNA expression level accuracy. A set of important genes were selected for qRT-PCR assessment. Some of them were from the highest-ranking potentially crucial pathway based on the pathway topology analysis and the others were the immediate downstream genes of the highest-ranking potentially crucial pathway. The same RNA samples as those isolated for microarrays at 3, 8, 16, 24, 32, 40, and 48 hours were used for RT-PCR. TATA box-binding protein (TBP) was used as the internal control gene. It was performed in triplicate for control and rCFES-treated WI-38 cells of the same time point.

3. Results and Discussion

3.1. Cytotoxic Effect of rCFES on WI-38 Cells. To understand the effect of recombinant CFP-10/ESAT-6 protein (rCFES) treatment on the survival of human lung fibroblasts, rCFES was purified by affinity chromatography. To ensure the treatment response was due to rCFES, a purity of higher than 95% was achieved as determined by densitometric scanning of the 12% SDS-PAGE (Figure 1). Then, WI-38 cells were treated with different concentrations of rCFES and incubated for 48 hours. Cell viability of WI-38 cells was determined by MTS assay. It was found that the survival rate of WI-38 cells decreased and the decrement varied with the different concentrations of rCFES. For example, only 62% WI-38 cells survived after 12.5 μM rCFES treatment (Figure 2), which suggested that rCFES influenced cell survival and consequently caused death in 48 hours. Based on these evidences, we inferred that rCFES could induce WI-38 cell death via cytotoxic process.

3.2. Identification of Important Genes. To identify the important genes involved in the rCFES-regulated mechanism of WI-38 cells, the microarray data for 0, 3, 8, 16, 24, 32, 40, and 48 hours were analyzed by SAM. Totally, 6542 important genes were identified with the following parameters: data type = one class time course, time summary method = signed
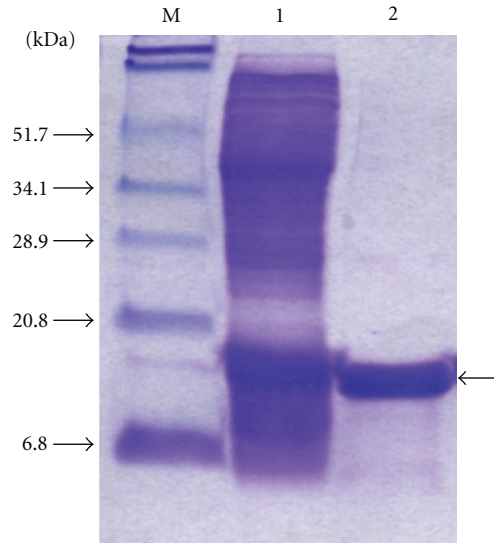

Figure 1: SDS-PAGE analysis of the purified rCFES. Lane M: Protein marker; Lane 1: IPTG-induced cell lysate; Lane 2: purified rCFES.
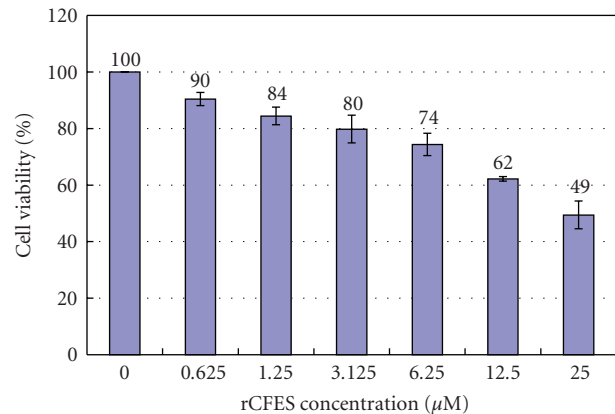

Figure 2: Effects of rCFES on WI-38 cell viability, in which the bars indicated the survival rates of WI-38 cells treated with different concentrations of rCFES. The data are expressed as mean values ($\pm$ standard deviation) of three experiments.

area, $\Delta = 0.511$ at time course, and false discovery rate = 4.74%. Out of these 6542 important genes, 2074 genes were upregulated and 4468 genes were downregulated with rCFES treatment.

3.3. Potentially Crucial Pathways in rCFES-Induced WI-38 Cells. To unravel the cytotoxic effect of rCFES on WI-38 cells, an integrative analysis approach was proposed in this study with the emphasis on identification of the potentially crucial pathways. While the high-throughput biological experiments, for example, microarray data, may provide hundreds or thousands of important genes revealing changes of expressions induced by rCFES, the information at the gene level may be too overwhelming to extract the most important regulatory effects. Moreover, the analysis results

are generally lack of the notion at the system level. To remedy these deficiencies, the proposed integrative analysis approach aimed at finding the pathways that played crucial roles in rCFES-induced WI-38 cells.

We integrated the significance of gene expressions and topological distribution of important genes to account for the regulatory effects at the gene and system levels. While important genes represented the regulatory effects of rCFES on individual genes, topological distribution of the important genes bore the notion of chaining or aggregating regulatory effects on genetic subnetworks of WI-38 cells. Suppose that there were two pathways containing the same number of important genes with similar significance, if the important genes distributed sporadically in one pathway and were highly connected in the other, the proposed approach preferred the latter to the former. This was because a cluster of connected genes implied a chaining or aggregating regulatory effect of rCFES on WI-38 cells, which deserved a further investigation.

With a control of the false discovery rate at 5%, 23 pathways as listed in Table 1 were identified to be the candidates of potentially crucial pathways in rCFES-induced WI-38 cells based on the hyper geometric statistic test. Ranking these 23 pathways according to the composite score, R_r, the second step of pathway topology analysis further suggested that the focal adhesion pathway is the potentially most crucial pathway in rCFES-induced WI-38 cells. Other than the focal adhesion pathway, the proposed integrative analysis approach also suggested that several metabolism pathways of WI-38 cells (as listed in Table 1) might play crucial role in the response to the treatment of rCFES. While further investigation would be required to validate this finding, it was partially supported by some previous studies, in which antibacterial genes of the immune cells infected with MTB were found to be involved in metabolism [23, 24].

To characterize the functional significance of the focal adhesion pathway, the topological and functional properties of the important genes involved in this pathway were elaborately analyzed. With the false discovery rate of 4.74%, 65 important genes, including 7 upregulated genes and 58 downregulated genes (Table 2), were selected by SAM for the focal adhesion pathway annotated by the KEGG database. The root gene of the MD-cluster identified by the pathway topology analysis was ITGB1. The functional interpretations of these 65 important genes were annotated by GOstat [25] using biological process terms at P-value $< 2 * 10^{-20}$ and numbers of gene >16, the results of which were shown in Figure 3. The GO terms that best described the functional disorders associated with these 65 important genes were developmental process, cell motility, localization of cell, biological adhesion, and cell adhesion. Most gene functions of 65 important genes were downregulated by cytotoxic effect of rCFES. Combining the analysis results of gene ontology, pathway topology analysis and annotation of KEGG database, overall speaking, the biological process of WI-38 cells treated with rCFES involved such functions as cell proliferation, cell motility, cell survival, amino acid metabolism, and so on. In particular, many of them were related to cell survival.
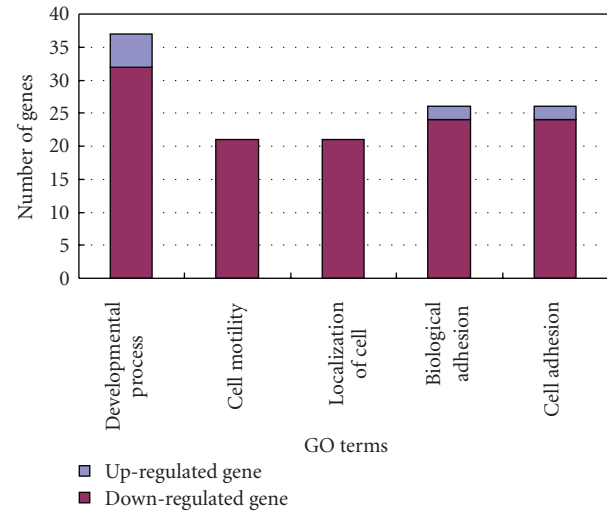

FIGURE 3: Functional analysis of the 65 important genes in the focal adhesion pathway using GOstat, in which five GO terms were identified to best describe the functional disorders associated with these 65 important genes, including developmental process, cell motility, localization of cell, biological adhesion, and cell adhesion.

To check if the proposed approach is also applicable to other microarray data, the same approach is used to analyze a set of publicly available human neutrophil microarray data. The analysis results may be found in Supplemental Data II (http://homepage.ntu.edu.tw/~d91548013/CFES_supp_data2.pdf).

3.4. Validation of Important Genes by qRT-PCR. Twenty-three pathways were identified as the potentially crucial pathways in rCFES-induced WI-38 cells and the focal adhesion pathway was considered as the most crucial one. These findings were based on the correctness of the gene expressions revealed by microarrays. To further corroborate the basis of the pathway topology analysis, the gene expressions of 5 important genes involved in the focal adhesion pathway were validated by qRT-PCR. These 5 genes were LAMA4, ITGB1, PIK3R3, BIRC3, and NFKBIA. According to gene annotations from KEGG databases, ITGB1 was the root gene of the focal adhesion pathway. LAMA4 was an upstream gene of ITGB1. PIK3R3 was a downstream gene of ITGB1. BIRC3 and NFKBIA were the downstream genes of PIK3R3. LAMA4 was also related to the ECM receptor interaction pathway. BIRC3 and NFKBIA were directly involved in the apoptosis pathway.

Figure 4 provided the fold changes measured by microarray data and qRT-PCR for these 5 important genes at each time point. The qRT-PCR results were generally in congruence with the microarray data for LAMA4, PIK3R3, BIRC3, and NFKBIA in terms of up- or downregulation (Pearson correlations: 0.66~0.82). More specifically, the first two genes, LAMA4 and PIK3R3, were downregulated by rCFES treatment and the last two genes, BIRC3, and NFKBIA, were upregulated by rCFES treatment. Nevertheless, the microarray data of ITGB1 was not in agreement with

Table 1: The values of the descriptors constituting the composite score, R_r, and the selected pathways sorted in the ascending order of R_r derived by the pathway topology analysis.

Pathway			Density		Average significance		MD-Cluster		MUD-Cluster		Composite score
Entry	Name	K	Q_d	Q_{pd}	Q_s	Q_{ps}	N_{cd}	Q_{cd}	N_{cu}	Q_{cu}	R_r
hsa04510	Focal adhesion	65	0.3351	0.2697	0.00130	0.00000039	11	0.00177	13	0.00200	46
hsa00450	Selenoamino acid metabolism	14	0.4118	0.5833	0.00119	0.00000365	6	0.00195	7	0.00221	49
hsa00230	Purine metabolism	54	0.3506	0.4230	0.00501	0.00002726	44	0.00566	51	0.00523	59
hsa00271	Methionine metabolism	11	0.6875	0.3158	0.00239	0.00001452	4	0.00044	7	0.00239	59.5
hsa00260	Glycine, serine and threonine metabolism	16	0.3556	0.1391	0.00168	0.00000594	6	0.00153	8	0.00082	61.5
hsa00330	Arginine and proline metabolism	21	0.3448	0.2075	0.00207	0.00001056	10	0.00223	14	0.00388	64
hsa04512	ECM-receptor interaction	31	0.3563	0.0909	0.00076	0.00000023	2	0.00010	5	0.00058	71.5
hsa00280	Valine, leucine and isoleucine degradation	19	0.3519	0.3913	0.00636	0.00002807	10	0.00657	14	0.00760	72
hsa00310	Lysine degradation	18	0.3462	0.4231	0.00453	0.00000841	5	0.01567	5	0.00035	82
hsa00100	Biosynthesis of steroids	8	0.4444	0.3902	0.00754	0.00006827	9	0.00769	9	0.00769	87
hsa04670	Leukocyte transendothelial migration	31	0.2650	0.0656	0.00135	0.00000028	2	0.00007	3	0.00030	87.5
hsa04910	Insulin signaling pathway	41	0.3037	0.1647	0.00713	0.00001028	8	0.00535	9	0.00583	89.5
hsa00380	Tryptophan metabolism	30	0.3488	0.2289	0.00850	0.00006979	6	0.00455	7	0.00543	100
hsa00130	Ubiquinone biosynthesis	7	0.4667	0.2500	0.00672	0.00003048	2	0.00552	2	0.00552	100
hsa00564	Glycerophospholipid metabolism	19	0.2468	0.2464	0.00692	0.00005076	7	0.00742	10	0.00786	103.5
hsa00020	Citrate cycle (TCA cycle)	14	0.4643	0.1121	0.00959	0.00006191	7	0.00758	7	0.00758	105
hsa00650	Butanoate metabolism	20	0.4545	0.2326	0.00713	0.00007358	6	0.00977	6	0.00977	108
hsa04514	Cell adhesion molecules (CAMs)	36	0.2727	0.1565	0.00305	0.00009319	2	0.00195	2	0.00195	112
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	13	0.2826	0.1200	0.00513	0.00012724	3	0.00875	3	0.00875	135.5
hsa00500	Starch and sucrose metabolism	23	0.2949	0.1250	0.00955	0.00018004	5	0.01366	6	0.01538	141.5
hsa00625	Tetrachloroethene degradation	4	0.4444	0.1250	0.02087	0.00043538	2	0.02087	2	0.02087	149
hsa00970	Aminoacyl-tRNA biosynthesis	12	0.3750	0.0476	0.00654	0.00077841	2	0.02790	2	0.02790	154
hsa00040	Pentose and glucuronate interconversions	6	0.2500	0.0588	0.00915	0.00050551	2	0.02248	2	0.02248	171

the readings of qRT-PCR (Pearson correlations <0), though ITGB1 was identified as the root gene in the focal adhesion pathway. One reasonable explanation for this inconsistency was that the microarray data were inherently noisy, and there was a false discovery rate associated with the microarray data analysis for ITGB1.

Among the confirmed genes, the downregulation of LAMA4 would affect cell survival via laminin-integrin interaction [26, 27] and the downregulation of PI3K would affect the survival signal transduction via an integrin/PI3K/Akt pathway [28]. Furthermore, the upregulation of baculoviral IAP repeat-containing 3 (BIRC3), known as an inhibitor of

TABLE 2: Gene names, gene descriptions, and significance analysis results of mapped genes in focal adhesion pathway obtained via SAM analysis, where "[ ]" showed upregulated (+) or downregulated (−) by rCFES treatment at q-value <5%.

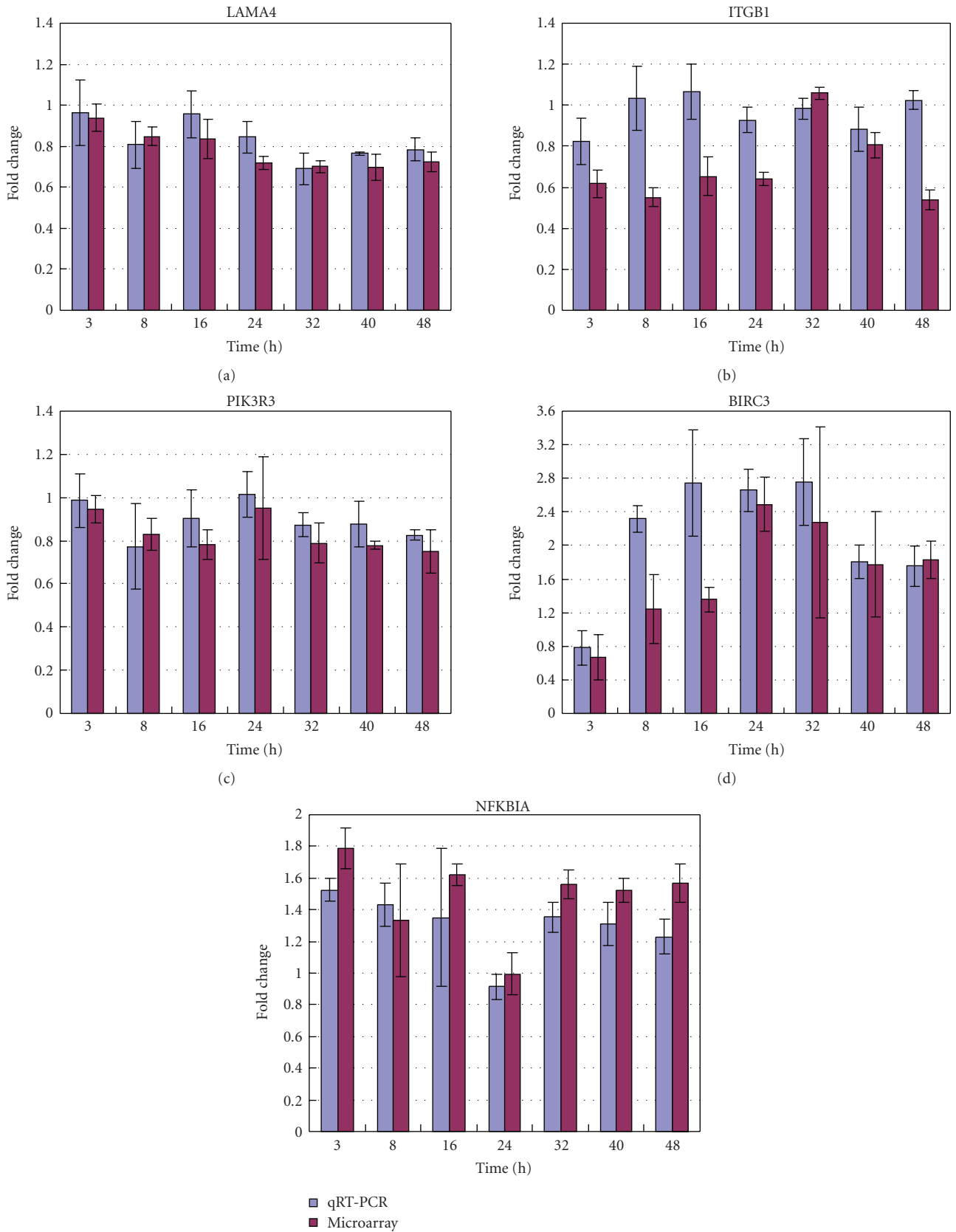| No.	Gene	Gene Description	q-value
1	SHC3	SHC transforming protein 3	0.0000 [+]
2	ACTN2	Actinin, alpha 2	0.0068 [+]
3	BIRC3	Baculoviral IAP repeat-containing 3	0.0211 [+]
4	SPP1	Secreted phosphoprotein 1	0.0211 [+]
5	AF086032	Protein phosphatase 1, regulatory (inhibitor) subunit 12A	0.0278 [+]
6	KDR	Kinase insert domain receptor	0.0278 [+]
7	DDI1	DDI1, DNA-damage inducible 1,homolog 1 (S. cerevisiae)	0.0278 [+]
8	ACTB	Actin, beta	0.0000 [−]
9	FLNA	Filamin A, alpha (actin binding protein 280)	0.0000 [−]
10	ACTG1	Actin, gamma 1	0.0000 [−]
11	CAV1	Caveolin 1, caveolae protein, 22kDa	0.0000 [−]
12	ITGB1	Integrin, beta 1	0.0000 [−]
13	ITGA6	Integrin, alpha 6	0.0000 [−]
14	COL1A2	Collagen, type I, alpha 2	0.0000 [−]
15	CAPN2	Calpain 2, (m/II) large subunit	0.0000 [−]
16	AKT1	V-akt murine thymoma viral oncogene homolog 1	0.0010 [−]
17	TNXB	Tenascin XB	0.0010 [−]
18	ITGA5	Integrin, alpha 5 (fibronectin receptor, alpha polypeptide)	0.0010 [−]
19	RAPGEF1	Rap guanine nucleotide exchange factor (GEF) 1	0.0014 [−]
20	ITGA8	Integrin, alpha 8	0.0014 [−]
21	ITGA4	Integrin, alpha 4	0.0014 [−]
22	PIK3R3	Phosphoinositide-3-kinase,regulatory subunit 3 (p55, gamma)	0.0014 [−]
23	FN1	Fibronectin 1	0.0014 [−]
24	LAMA4	Laminin, alpha 4	0.0016 [−]
25	THBS1	Thrombospondin 1	0.0016 [−]
26	IGF1R	Insulin-like growth factor 1 receptor	0.0016 [−]
27	COL1A1	Collagen, type I, alpha 1	0.0016 [−]
28	CCND1	Cyclin D1	0.0016 [−]
29	MET	Met proto-oncogene (hepatocyte growth factor receptor)	0.0016 [−]
30	ACTN4	Actinin, alpha 4	0.0019 [−]
31	AK025363	Parvin, alpha	0.0019 [−]
32	AKT2	V-akt murine thymoma viral oncogene homolog 2	0.0019 [−]
33	CAV2	Caveolin 2	0.0027 [−]
34	VCL	Vinculin	0.0027 [−]
35	ZYX	Zyxin	0.0027 [−]
36	TNC	Tenascin C (hexabrachion)	0.0027 [−]
37	ACTN3	Actinin, alpha 3	0.0028 [−]
38	VASP	Vasodilator-stimulated phosphoprotein	0.0028 [−]
39	COL5A1	Collagen, type V, alpha 1	0.0028 [−]
40	VEGFC	Vascular endothelial growth factor C	0.0039 [−]
41	FAK	Focal adhesion kinase	0.0047 [−]
42	PARVB	Parvin, beta	0.0047 [−]
43	COL4A6	Collagen, type IV, alpha 6	0.0047 [−]
44	LAMA5	Laminin, alpha 5	0.0068 [−]
45	MAPK1	Mitogen-activated protein kinase 1	0.0088 [−]
46	AK025363	Parvin, alpha	0.0019 [−]
47	THBS4	Thrombospondin 4	0.0116 [−]
48	ERBB2	V-erb-b2 erythroblastic leukemiaviral oncogene homolog 2	0.0157 [−]
49	AF207599	Homo sapiens pRb-interacting protein RbBP-36 mRNA	0.0157 [−]
50	DOCK1	Dedicator of cytokinesis 1	0.0157 [−]
51	COL5A2	Collagen, type V, alpha 2	0.0211 [−]

FIGURE 4: Validation of microarray data by qRT-PCR analysis for 5 genes: (a) LAMA4, (b) ITGB1, (c) PIK3R3, (d) BIRC3, and (e) NFKBIA. The data are expressed as mean values (± standard deviation) of three experiments.

TABLE 2: Continued.

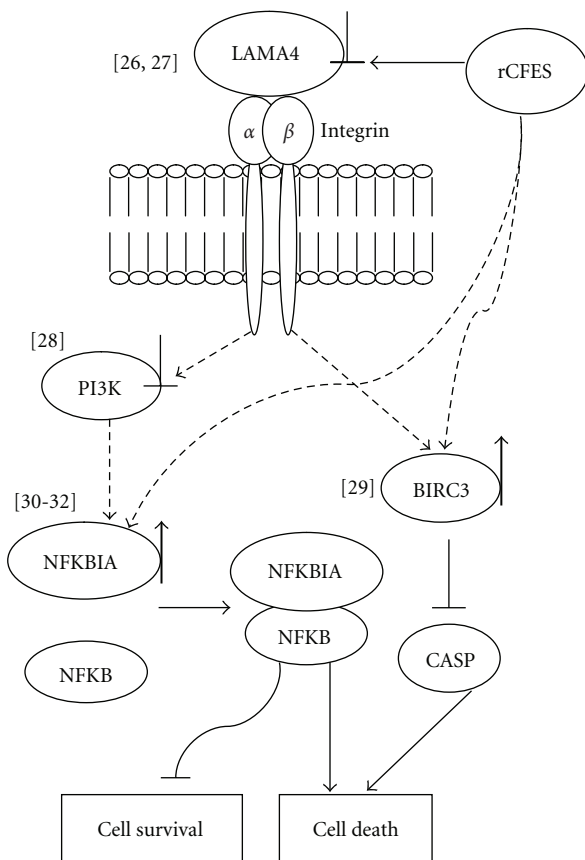| No.	Gene	Gene Description	q-value
52	PIK3R5	Phosphoinositide-3-kinase, regulatory subunit 5, p101	0.0211 [−]
53	ITGAV	Integrin, alpha V	0.0211 [−]
54	LAMA2	Laminin, alpha 2	0.0278 [−]
55	ARHGAP5	Rho GTPase activating protein 5	0.0278 [−]
56	VAV2	Vav 2 oncogene	0.0278 [−]
57	CAV3	Caveolin 3	0.0278 [−]
58	BCAR1	Breast cancer anti-estrogen resistance 1	0.0369 [−]
59	EGFR	Epidermal growth factor receptor	0.0369 [−]
60	MAPK9	Mitogen-activated protein kinase 9	0.0369 [−]
61	AA599881	Platelet-derived growth factor receptor, alpha polypeptide	0.0474 [−]
62	THBS1	Thrombospondin 1	0.0016 [−]
63	GRB2	Growth factor receptor-bound protein 2	0.0474 [−]
64	LAMA3	Laminin, alpha 3	0.0474 [−]
65	CRKL	V-crk sarcoma virus CT10 oncogene homolog (avian)-like	0.0474 [−]

FIGURE 5: Inferred model for the rCFES-regulated focal adhesion pathway in WI-38 cells. LAMA4, PIK3R3, BIRC3, and NFKBIA identified in this study were included in the rCFES-regulated model in WI-38 cells. Upregulated genes were marked by "↑" and downregulated genes by "⊥". Dotted arrows were links requiring further confirmation and the numbers in "[ ]" were the supportive previous works listed in the reference.

caspase family [29], would inhibit the activation of caspase genes and the upregulation of NFKBIA would induce cell

apoptosis via interfering with NF-κB activation [30–32]. These results supported that rCFES could influence WI-38 cell survivability via the regulation of focal adhesion pathway at least. Based on these results, a working model was inferred for the rCFES-regulated pathway in WI-38 cells in Figure 5. Although this model needed further verification, it was believed that these results will aid the investigation of the CFES-regulated mechanism in human lung fibroblasts.

Acknowledgments

References

[1] M. C. Raviglione, "The TB epidemic from 1992 to 2002," *Tuberculosis*, vol. 83, no. 1–3, pp. 4–14, 2003.

[2] W. H. Boom, D. H. Canaday, S. A. Fulton, A. J. Gehring, R. E. Rojas, and M. Torres, "Human immunity to *M. tuberculosis*: T cell subsets and antigen processing," *Tuberculosis*, vol. 83, no. 1–3, pp. 98–106, 2003.

[3] J. L. Flynn and J. Chan, "Immunology of tuberculosis," *Annual Review of Immunology*, vol. 19, pp. 93–129, 2001.

[4] C. M. O'Kane, J. J. Boyle, D. E. Horncastle, P. T. Elkington, and J. S. Friedland, "Monocyte-dependent fibroblast CXCL8 secretion occurs in tuberculosis and limits survival of mycobacteria within macrophages," *Journal of Immunology*, vol. 178, no. 6, pp. 3767–3776, 2007.

[5] C. M. O'Kane, P. T. Elkington, and J. S. Friedland, "Monocyte-dependent oncostatin M and TNF-α synergize to stimulate unopposed matrix metalloproteinase-1/3 secretion from human lung fibroblasts in tuberculosis," *European Journal of Immunology*, vol. 38, no. 5, pp. 1321–1330, 2008.

[6] T. Hsu, S. M. Hingley-Wilson, B. Chen, et al., "The primary mechanism of attenuation of bacillus Calmette-Guérin is a loss of secreted lytic function required for invasion of lung interstitial tissue," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 21, pp. 12420–12425, 2003.

[7] F. Tekaia, S. V. Gordon, T. Garnier, R. Brosch, B. G. Barrell, and S. T. Cole, "Analysis of the proteome of *Mycobacterium tuberculosis* in silico," *Tubercle and Lung Disease*, vol. 79, no. 6, pp. 329–342, 1999.

[8] K. M. Guinn, M. J. Hickey, S. K. Mathur, et al., "Individual RD1 -region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*," *Molecular Microbiology*, vol. 51, no. 2, pp. 359–370, 2004.

[9] L.-Y. Gao, S. Guo, B. McLaughlin, H. Morisaki, J. N. Engel, and E. J. Brown, "A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion," *Molecular Microbiology*, vol. 53, no. 6, pp. 1677–1693, 2004.

[10] N. van der Wel, D. Hava, D. Houben, et al., "*M. tuberculosis* and *M. leprae* translocate from the phagolysosome to the cytosol in myeloid cells," *Cell*, vol. 129, no. 7, pp. 1287–1298, 2007.

[11] S. C. Derrick and S. L. Morris, "The ESAT6 protein of *Mycobacterium tuberculosis* induces apoptosis of macrophages by activating caspase expression," *Cellular Microbiology*, vol. 9, no. 6, pp. 1547–1555, 2007.

[12] P. S. Renshaw, P. Panagiotidou, A. Whelan, et al., "Conclusive evidence that the major T-cell antigens of the *Mycobacterium tuberculosis* complex ESAT-6 and CFP-10 form a tight, 1:1 complex and characterization of the structural properties of ESAT-6, CFP-10, and the ESAT-6·CFP-10 complex. Implications for pathogenesis and virulence," *Journal of Biological Chemistry*, vol. 277, no. 24, pp. 21598–21603, 2002.

[13] A. K. Meher, R. K. Lella, C. Sharma, and A. Arora, "Analysis of complex formation and immune response of CFP-10 and ESAT-6 mutants," *Vaccine*, vol. 25, no. 32, pp. 6098–6106, 2007.

[14] C. Abramo, K. E. Meijgaarden, D. Garcia, et al., "Monokine induced by interferon gamma and IFN-γ response to a fusion protein of *Mycobacterium tuberculosis* ESAT-6 and CFP-10 in Brazilian tuberculosis patients," *Microbes and Infection*, vol. 8, no. 1, pp. 45–51, 2006.

[15] V. G. Tusher, R. Tibshirani, and G. Chu, "Significance analysis of microarrays applied to the ionizing radiation response," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 9, pp. 5116–5121, 2001.

[16] N. Jain, J. Thatte, T. Braciale, K. Ley, M. O'Connell, and J. K. Lee, "Local-pooled-error test for identifying differentially expressed genes with a small number of replicated microarrays," *Bioinformatics*, vol. 19, no. 15, pp. 1945–1951, 2003.

[17] P. Baldi and A. D. Long, "A Bayesian framework for the analysis of microarray expression data: regularized t-test and statistical inferences of gene changes," *Bioinformatics*, vol. 17, no. 6, pp. 509–519, 2001.

[18] K. D. Dahlquist, N. Salomonis, K. Vranizan, S. C. Lawlor, and B. R. Conklin, "GenMAPP, a new tool for viewing and analyzing microarray data on biological pathways," *Nature Genetics*, vol. 31, no. 1, pp. 19–20, 2002.

[19] M. Hewett, D. E. Oliver, D. L. Rubin, et al., "PharmGKB: the pharmacogenetics knowledge base," *Nucleic Acids Research*, vol. 30, no. 1, pp. 163–165, 2002.

[20] M. Kanehisa and S. Goto, "KEGG: Kyoto encyclopedia of genes and genomes," *Nucleic Acids Research*, vol. 28, no. 1, pp. 27–30, 2000.

[21] H.-J. Chung, C. H. Park, M. R. Han, et al., "ArrayXPath II: mapping and visualizing microarray gene-expression data with biomedical ontologies and integrated biological pathway resources using Scalable Vector Graphics," *Nucleic Acids Research*, vol. 33, pp. W621–W626, 2005.

[22] J. D. Storey and R. Tibshirani, "Statistical significance for genomewide studies," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 16, pp. 9440–9445, 2003.

[23] D. Chaussabel, R. T. Semnani, M. A. McDowell, D. Sacks, A. Sher, and T. B. Nutman, "Unique gene expression profiles of human macrophages and dendritic cells to phylogenetically distinct parasites," *Blood*, vol. 102, no. 2, pp. 672–681, 2003.

[24] S. Ehrt, D. Schnappinger, S. Bekiranov, et al., "Reprogramming of the macrophage transcriptome in response to interferon-γ and *Mycobacterium tuberculosis*: signaling roles of nitric oxide synthase-2 and phagocyte oxidase," *Journal of Experimental Medicine*, vol. 194, no. 8, pp. 1123–1140, 2001.

[25] T. Beissbarth and T. P. Speed, "GOstat: find statistically overrepresented gene ontologies with a group of genes," *Bioinformatics*, vol. 20, no. 9, pp. 1464–1465, 2004.

[26] K. C. DeHahn, M. Gonzales, A. M. Gonzalez, et al., "The α4 laminin subunit regulates endothelial cell survival," *Experimental Cell Research*, vol. 294, no. 1, pp. 281–289, 2004.

[27] A. M. Gonzalez, M. Gonzales, G. S. Herron, et al., "Complex interactions between the laminin α4 subunit and integrins regulate endothelial cell behavior in vitro and angiogenesis in vivo," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 25, pp. 16075–16080, 2002.

[28] H. Xia, R. S. Nho, J. Kahm, J. Kleidon, and C. A. Henke, "Focal adhesion kinase is upstream of phosphatidylinositol 3-kinase/Akt in regulating fibroblast survival in response to contraction of type I collagen matrices via a β1 integrin viability signaling pathway," *Journal of Biological Chemistry*, vol. 279, no. 31, pp. 33024–33034, 2004.

[29] Q. L. Deveraux and J. C. Reed, "IAP family proteins—suppressors of apoptosis," *Genes and Development*, vol. 13, no. 3, pp. 239–252, 1999.

[30] I. Jeremias, C. Kupatt, B. Baumann, I. Herr, T. Wirth, and K. M. Debatin, "Inhibition of nuclear factor κB activation attenuates apoptosis resistance in lymphoid cells," *Blood*, vol. 91, no. 12, pp. 4624–4631, 1998.

[31] F. Arenzana-Seisdedos, J. Thompson, M. S. Rodriguez, F. Bachelerie, D. Thomas, and R. T. Hay, "Inducible nuclear expression of newly synthesized IκBα negatively regulates DNA-binding and transcriptional activities of NF-κB," *Molecular and Cellular Biology*, vol. 15, no. 5, pp. 2689–2696, 1995.

[32] F. Arenzana-Seisdedos, P. Turpin, M. Rodriguez, et al., "Nuclear localization of IκBα promotes active transport of NF-κB from the nucleus to the cytoplasm," *Journal of Cell Science*, vol. 110, no. 3, pp. 369–378, 1997.