Toxicology and Applied Pharmacology 272 (2013) 476–489

Contents lists available at ScienceDirect

Toxicology and Applied Pharmacology

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

System-based identification of toxicity pathways associated with multi-walled carbon nanotube-induced pathological responses

Brandi N. Snyder-Talkington [a,1], Julian Dymacek [b,c,1], Dale W. Porter [a], Michael G. Wolfarth [a], Robert R. Mercer [a], Maricica Pacurari [c], James Denvir [d], Vincent Castranova [a], Yong Qian [a,*], Nancy L. Guo [c,**]

[a] Pathology and Physiology Research Branch, Health Effects Laboratory Division, National Institute for Occupational Safety and Health, Morgantown, WV 26505, USA
[b] Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506-6070, USA
[c] Mary Babb Randolph Cancer Center, West Virginia University, Morgantown, WV 26506-9300, USA
[d] Department of Biochemistry and Microbiology, Marshall University, Huntington, WV 25755, USA

ARTICLE INFO

Article history:
Received 29 April 2013
Revised 24 June 2013
Accepted 25 June 2013
Available online 8 July 2013

Keywords:
Multi-walled carbon nanotubes
Signaling pathways
Computational toxicology
In vivo studies
In vitro studies

ABSTRACT

The fibrous shape and biopersistence of multi-walled carbon nanotubes (MWCNT) have raised concern over their potential toxicity after pulmonary exposure. As *in vivo* exposure to MWCNT produced a transient inflammatory and progressive fibrotic response, this study sought to identify significant biological processes associated with lung inflammation and fibrosis pathology data, based upon whole genome mRNA expression, bronchoaveolar lavage scores, and morphometric analysis from C57BL/6J mice exposed by pharyngeal aspiration to 0, 10, 20, 40, or 80 μg MWCNT at 1, 7, 28, or 56 days post-exposure. Using a novel computational model employing non-negative matrix factorization and Monte Carlo Markov Chain simulation, significant biological processes with expression similar to MWCNT-induced lung inflammation and fibrosis pathology data in mice were identified. A subset of genes in these processes was determined to be functionally related to either fibrosis or inflammation by Ingenuity Pathway Analysis and was used to determine potential significant signaling cascades. Two genes determined to be functionally related to inflammation and fibrosis, vascular endothelial growth factor A (*vegfa*) and C-C motif chemokine 2 (*ccl2*), were confirmed by *in vitro* studies of mRNA and protein expression in small airway epithelial cells exposed to MWCNT as concordant with *in vivo* expression. This study identified that the novel computational model was sufficient to determine biological processes strongly associated with the pathology of lung inflammation and fibrosis and could identify potential toxicity signaling pathways and mechanisms of MWCNT exposure which could be used for future animal studies to support human risk assessment and intervention efforts.

Introduction

Nanotechnology is an emerging discipline in both industrial and medical fields, which necessitates the development of nanotoxicology to determine the biological effects of occupational and commercial nanoparticle exposures (Oberdorster et al., 2005). Multi-walled carbon nanotubes (MWCNT) are fibrous nanoparticles consisting of multiple concentric cylindrical carbon tubes that are appealing for both industrial and medical purposes due to their efficient electronic conductivity, great strength, and strong capillary forces, while maintaining a small size, light weight, high surface area to mass ratio, and low density (Ajayan, 1999; Castranova, 2011; Iijima, 1991). The physical attributes of MWCNT, while useful from an engineering standpoint, make them easily aerosolized and a potential inhalation hazard during synthesis, product use, and disposal. *In vitro* studies of MWCNT exposure determined toxicity to both lung epithelial and microvascular endothelial cells with increases in reactive oxygen species (ROS) production, NF-κB signaling, cytokine release, cytoskeletal reorganization, and endothelial cell permeability (He et al., 2011; Pacurari et al., 2012; Srivastava et al., 2011; Walker et al., 2009; Ye et al., 2009). Mouse and rat *in vivo* studies determined that MWCNT can reach the alveolar region of the lung after pharyngeal aspiration and inhalation, respectively, and induce a transient inflammatory reaction followed by a progressive fibrotic response (Mercer et al., 2010, 2011; Muller et al., 2005).

Although chronic inflammation has been suggested as the underlying mechanism governing the progression to fibrosis, this does not appear to hold true for MWCNT as the initial inflammatory response

* Correspondence to: Y. Qian, Pathology and Physiology Research Branch, Health Effects Laboratory Division, National Institute for Occupational Safety and Health, 1095 Willowdale Road, Morgantown, WV 26505-2888. Fax: +1 304 285 5938.
** Correspondence to: N.L. Guo, Mary Babb Randolph Cancer Center and Department of Occupational and Environmental Health Sciences, West Virginia University, Morgantown, WV 26506-9300, USA. Fax: +1 304 293 4667.
E-mail addresses: yaq2@cdc.gov (Y. Qian), lguo@hsc.wvu.edu (N.L. Guo).
[1] Both authors contributed equally to this manuscript.

0041-008X
http://dx.doi.org/10.1016/j.taap.2013.06.026

to MWCNT diminishes before the progressive fibrotic response begins (Mercer et al., 2010, 2011; Porter et al., 2010). Therefore, it is essential to uncover the significant biological processes directing MWCNT-induced inflammation and fibrosis so as to determine potential outcomes and hallmarks of exposure. We hypothesize that the identification of transcription-related biological processes and pathways, which match the patterns of BAL quantification (Porter et al., 2010) for inflammatory pathology and morphometric scoring of collagen (Mercer et al., 2011) for fibrosis in MWCNT-treated mice, could identify critical toxicity pathways and potential mechanisms of MWCNT-induced lung inflammation and fibrosis for early identification and intervention.

Recently, our group conducted an *in vivo* dose–response time-course study of MWCNT exposure in C57BL/6J mice to determine the ability of MWCNT to induce pulmonary inflammation, damage, and fibrosis (Porter et al., 2010). Mice were exposed to 0, 10, 20, 40, or 80 μg of MWCNT by pharyngeal aspiration with endpoints monitored at 1, 7, 28, and 56 days post-exposure (Porter et al., 2010). The results indicated that a transient inflammatory response occurred 1 day post-exposure with peak activity 7 days post-exposure. A fibrotic response was noted 28 days post-exposure, which progressed through 56 days post-exposure (Porter et al., 2010). Nevertheless, MWCNT-induced toxicity pathways and mechanisms underlying these observed *in vivo* pathological responses remain unknown. We hypothesize that systematic analyses of gene expression profiles and pathological data could identify transcription-related biological processes correlated with the observed pathological patterns of lung inflammation and fibrosis, which could reveal MWCNT-induced toxicity pathways and pathogenesis. The current study sought to use a novel computational system to identify transcription-related biological processes and pathways associated with these MWCNT-induced pathology responses in a comprehensive systematic evaluation. A novel computational model, previously reported by Dymacek and Guo (2011) was applied to genome-wide mRNA expression profiles and pathological analysis of mouse lungs taken at these respective time points so as to determine biological processes significantly correlated with inflammation (bronchoalveolar lavage fluid [BAL] score (Porter et al., 2010)) or fibrosis (morphometric analysis of alveolar interstitial fibrosis (Mercer et al., 2011)). These biological processes were then analyzed through Ingenuity Pathway Analysis (IPA) to determine gene subsets functionally related to inflammation or fibrosis. *In vitro* gene and protein expression data of two genes functionally related to inflammation and fibrosis, vascular endothelial growth factor A (*vegfa*) and C-C motif chemokine 2 (*ccl2*), were validated through cell culture studies.

This study determined that a novel computational model was sufficient to identify transcription-related biological processes strongly associated with lung inflammation BAL scores and fibrosis morphometric analysis. Potential toxicity signaling pathways of MWCNT exposure were determined and validated *in vitro*. The use of these toxicogenomic data and *in vivo* animal model-based gene expression profiling integrated with *in vitro* verification may allow for successful toxicity profiling of MWCNT as well as the identification of potential signaling pathways involved in the etiology of MWCNT-induced injury.

Materials and methods

MWCNT. MWCNT used in both mouse and cell studies were obtained from Mitsui & Company (MWCNT-7, lot #05072001K28) and have been previously characterized (Porter et al., 2010). Briefly, the bulk MWCNT exhibit a distinctive crystalline structure with the number of walls ranging from 20 to 50 walls. Overall, MWCNT trace metal contamination was 0.78%, including sodium (0.41%) and iron (0.32%) with no other trace metal contamination over 0.02%. Transmission electron microscopy (TEM) micrographs of MWCNT dispersed in dispersion medium (DM) demonstrated that DM promotes significant dispersion of MWCNT. The quantitative analysis of TEM micrographs revealed that the median length of the MWCNT sample was 3.86 μm (GSD 1.94) and the count mean width was 49 ± 13.4 (SD) nm. The zeta potential of the MWCNT in the DM was determined to be −11 mV.

Animals. Animal studies were performed as previously described (Porter et al., 2010). Briefly, male C57BL/6J mice (7 weeks old) were obtained from Jackson Laboratories (Bar Harbor, ME). Individual mice were housed one per cage in polycarbonate isolator ventilated cages and provided HEPA-filtered air with fluorescent lighting from 0700 to 1900 h. Autoclaved Alpha-Dri virgin cellulose chips and hardwood Beta-chips were used as bedding. Mice were monitored to be free of endogenous viral pathogens, parasites, mycoplasms, Helicobacter, and CAR Bacillus. Mice were maintained on Harlan Teklad Rodent Diet 7913 (Indianapolis, IN) and tap water was provided *ad libitum*. Animals were allowed to acclimate for at least 5 days before use. All animals in this study were housed in an AAALAC-accredited, specific pathogen-free, and environmentally controlled facility. All animal studies and procedures were approved by the National Institute for Occupational Safety and Health ACUC.

MWCNT pharyngeal aspiration exposure. Suspensions of MWCNT were prepared in DM and administered as previously described (Porter et al., 2008, 2010). In brief, each treatment group consisted of 8 mice, which were anesthetized with isoflurane (Abbott Laboratories, North Chicago, IL). When fully anesthetized, the mouse was positioned with its back against a slant board and suspended by the incisor teeth using a rubber band. The mouth was opened and the tongue gently pulled aside from the oral cavity. A 50 μl aliquot of sample was pipetted at the base of the tongue, and the tongue was restrained until at least 2 deep breaths were completed (but not for longer than 15 s). Following release of the tongue, the mouse was gently lifted off the board, placed on its left side, and monitored for recovery from anesthesia. Mice received either DM (vehicle control), or 10, 20, 40 or 80 μg MWCNT suspended in DM (Porter et al., 2010).

Tissue RNA extraction. Total RNA was extracted from frozen mouse lung tissue samples (−80 °C) in RNAlater using a RNeasy Fibrous Tissue Mini Kit according to manufacturer's protocol (Qiagen, USA) as previously described (Pacurari et al., 2011). Total RNA was eluted in RNase-free water and stored at −80 °C until further analysis. The quality and concentration of each RNA sample were determined using a NanoDrop-1000 Spectrophotometer (NanoDrop Tech, Germany).

Microarray expression profiling. Extracted RNA was analyzed for expression profiling using Agilent Mouse Whole Genome Arrays (Agilent, Santa Clara, CA). A universal reference design was employed using Stratagene Universal Mouse Reference RNA — Cat. No. 740100 (Agilent) as the reference RNA. Total RNA quality was determined on an Agilent 2100 Bioanalyzer with all samples having RNA integrity numbers (RIN) greater than 8. Total RNA (250 ng) was used for labeling, using the QuickAmp labeling kit (Agilent). RNA extracted from each mouse was labeled with cyanine (Cy)-3-CTP (PerkinElmer, Waltham, MA) and reference RNA with (Cy)-5-CTP. Following purification of labeled cRNAs, 825 ng of Cy3- and Cy5-labeled cRNAs were combined and hybridized for 17 h at 65 °C in an Agilent hybridization oven. Microarrays were then washed and scanned, using an Agilent DNA Microarray Scanner.

Pathological datasets. Inflammatory datasets were obtained by analysis of BAL fluid taken from MWCNT-exposed mice at 1, 7, 28 and 56 days post-exposure as previously described (Porter et al., 2010). Mice were euthanized with an i.p. injection of sodium pentobarbital (>100 mg/kg body weight) followed by exsanguination. A tracheal cannula was inserted and BAL was performed through the cannula using ice cold Ca^{2+} and Mg^{2+}-free phosphate buffered saline, pH 7.4,

supplemented with 5.5 mM D-glucose (PBS). The first lavage (0.6 ml) was kept separate from the rest of the lavage fluid. Subsequent lavages, each with 1 ml of PBS, were performed until a total of 4 ml of lavage fluid was collected. BAL cells were isolated by centrifugation (650 ×*g*, 5 min, 4 °C). An aliquot of the acellular supernatant from the first BAL (BAL fluid) was decanted and transferred to tubes for analysis of lactate dehydrogenase (LDH) and albumin. The acellular supernatants from the remaining lavage samples were decanted and discarded. BAL cells isolated from the first and subsequent lavages for the same mouse were pooled after resuspension in PBS, centrifuged a second time (650 ×*g*, 5 min, 4 °C), and the supernatant decanted and discarded. The BAL cell pellet was then resuspended in PBS and placed on ice. Total BAL cell counts were obtained using a Coulter Multisizer 3 (Coulter Electronics, Hialeah, FL) and cytospin preparations of the BAL cells were made using a cytocentrifuge (Shandon Elliot Cytocentrifuge, London). The cytospin preparations were stained with modified Wright-Giemsa stain and cell differentials were determined by light microscopy.

Fibrosis datasets were obtained by morphometric analysis of Sirius Red staining for connective tissue in MWCNT-exposed mice at 1, 7, 28 and 56 days post-exposure as previously described (Mercer et al., 2011). Briefly, mice were euthanized by an overdose of pentobarbital (>100 mg/kg body weight, i.p.) followed by transection of the abdominal aorta to provide exsanguination. To accomplish lung fixation, the trachea was cannulated and the lungs removed from the chest cavity. The lungs were then inflated with 1 ml of 10% neutral buffered formalin over a 1 minute period and the trachea tied off. After 4 to 5 h, the lungs were trimmed and processed overnight in a tissue processor. For each animal, the left lung lobe was placed in the embedding carrier with a consistent apex to base orientation and embedded in paraffin. For morphometric studies, paraffin sections of the left lung (5 μm thick) were cut. A new region of the disposable knife blade was used to section each block and the water bath was changed frequently in order to prevent potential cross-contamination that might result from MWCNT passage on the knife between sections. The sections were then deparaffinized and rehydrated with a xylene–alcohol series to distilled water. To enhance the contrast between tissue and MWCNT, lung sections were stained with Sirius Red (Junqueira et al., 1979). Sirius Red staining consisted of immersion of the slides in 0.1% Picrosirius solution (100 mg of Sirius Red F3BA in 100 ml of saturated aqueous picric acid, pH 2) for 1–2 h followed by washing for 1 min in 0.01 N HCl. Sections were then briefly counterstained in freshly filtered Mayer's hematoxylin for 2 min, dehydrated, and mounted on a slide with a coverslip. Quantitative morphometric methods were used to measure the average thickness of Sirius Red positive connective tissue fibers in the alveolar regions. Volume and surface density were measured using standard morphometric analyses (Weibel, 1980a,1980b). This consisted of basic point and intercept counting. Volume density was determined from counting the number of points over all tissues in the alveolar regions and points over Sirius Red positive connective tissue. Surface density of the alveolar wall was determined from intercepts between a line overlay and the alveolar wall. These point and intercept counts were made using a 121-point/11-line overlay graticule (12.5 mm square with 100 divisions), at 100× magnification, taken at six locations equally spaced across each section (one section per animal). This process was repeated twice for each animal. In order to limit the measurements to alveolar parenchyma, areas containing airways or blood vessels greater than 25 mm in diameter were excluded from the analysis. Average thickness of the Sirius Red positive connective tissue fibers of the alveolar wall was computed from two times the ratio of volume density of point to the surface density of the alveolar wall. The collagen fiber content of granulomatous lesions in the airspaces was assessed by a separate tabulation of points over Sirius Red positive connective tissues in granulomas and expressed as a percentage of total alveolar collagen. Mean linear intercept, a measure of the average size of the alveolar/alveolar duct airspaces in the alveolar region, was computed from the ratio of volume density to surface density (Weibel, 1980a).

Microarray data preprocessing and filtering. Data were exported from the Agilent DNA Microarray Scanner using Feature Extraction v10 as tab-delimited text files after background subtraction, log transformation, and lowess normalization and reported as log or relative expression of the sample compared to the universal reference. Data were read from each file into R using a custom script. For each array, values for control spots, spots which were saturated on either channel, and spots which were not well above background on at least one channel were considered unreliable and/or uninformative and were replaced by "NA". Values were collated into a single table, and probes for which fewer than 10 present values were available were removed. For probes spotted multiple times on the array, values were averaged across replicate probes. The resulting table is available as a series matrix in the NCBI Gene Expression Omnibus repository with accession number GSE29042 (Guo et al., 2012). A web-interface (http://www.mwcnttranscriptome.org) was developed to visualize the expression pattern of every gene in the whole genome in each MWCNT treatment condition.

Computational system. The computational system (Fig. 1B) was divided into four main components: a preprocessing component, the Pattern Finding component, the Coefficient Expander (CE) component, and the Functional Process Evaluation (FPE) component. Source code for the computational system can be found at http://sourceforge.net/projects/megpath.

First, the preprocessing step was used to identify probes with significant changes in expression. Missing data were imputed using the K-means nearest neighbor algorithm as implemented by the *impute.knn* function in the impute R package from Bioconductor (Seattle, WA). Using the Bioconductor package, a set of differentially expressed genes for each dose and time point was identified by performing a two-class unpaired Significance Analysis of Microarrays (SAM) between the treated samples and the dose zero samples from the corresponding time point. A threshold delta value was chosen to produce a false discovery rate of 1% using the *find Delta* function from the same package. The list of probes called as significant was subsequently filtered by restricting those probes which were at least 1.5 fold up- or down-regulated. Fold changes were computed from the data before imputation of missing values.

Additionally, a linear model was fit to the data, modeling the log expression of each gene as a function of time, dose, and the interaction of time with dose. The t-statistic associated with the dose and interaction parameters following the SAM algorithm was moderated and a threshold set to control for a false discovery rate of 0.1%, thus generating a list of genes whose expression values were significantly dependent on dose and a list of gene whose expression values were significantly dependent on dose in a time-dependent fashion. The combined list of probes was described by Guo et al. (2012) and was used by the Pattern Finding component in the current study.

Second, the Pattern Finding component was based on a Non-negative Matrix Factorization algorithm. This algorithm attempted to find a set of non-orthogonal basis vectors (patterns), which could be linearly combined to reconstruct the original probe expression data. In addition to finding the patterns, the algorithm also found coefficients relating each probe to each pattern. These coefficients could be used to describe how closely a probe matches a pattern. Most importantly, the Pattern Finding algorithm allowed for a probe to be associated with multiple patterns. In this way, patterns could be thought of as functions, such as fibrosis or inflammation, hence probes may be involved in multiple functions. The Pattern Finding algorithm worked as a Monte Carlo Markov chain with each location in the coefficient and pattern matrices having an associated probability

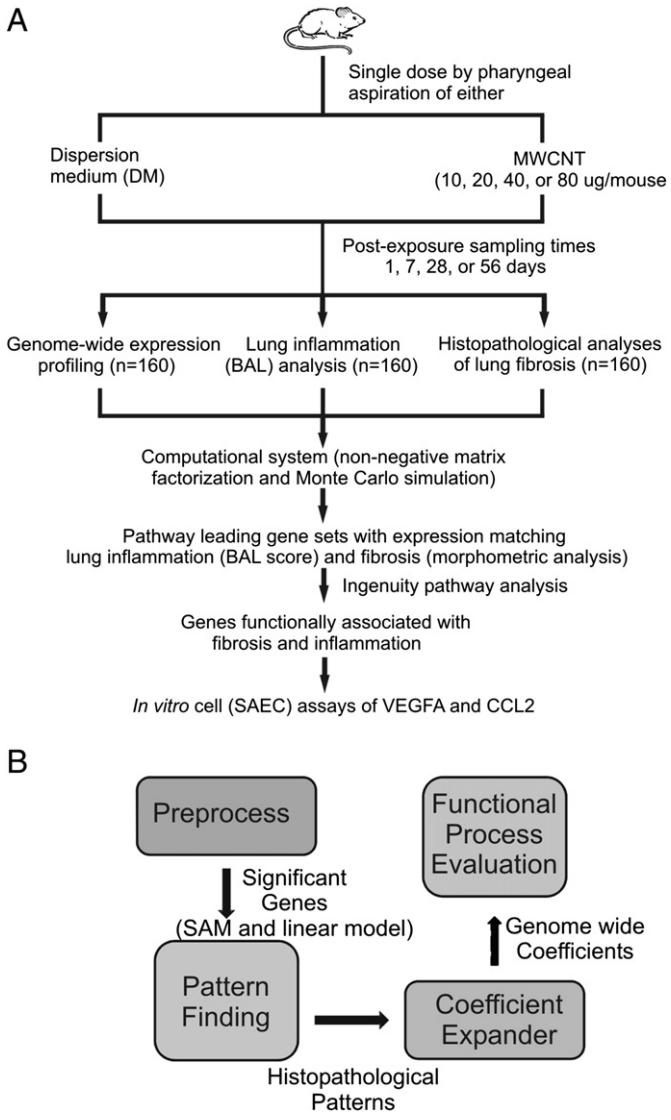

Fig. 1. (A) Schematic overview of MWCNT exposure, lung harvest, gene expression and histopathological analyses, computational system analysis, IPA analysis, and *in vitro* verification. (B) Overview of the four steps in the computational system. Step 1: Preprocessing to identify significantly changing genes and identify potential interesting genes. Step 2: Find patterns and coefficients from the interesting genes to reconstruct the gene expression data. Step 3: Find coefficients for the entire genome. Step 4: Using the patterns, coefficients, and pathways/functions, identify significant pathways.

density function. Quantitative histopathology data were used as one pattern.

The third step was to apply the Coefficient Expander component. This step attempted, through the use of simulated annealing, to find optimal coefficients for each probe in the genome-wide microarray data after preprocessing from the patterns found in the Pattern Finding step, and, therefore, enabled the probe's expression to be reconstructed from the patterns with minimal error.

The final step was to calculate the Functional Process Evaluation (FPE) score for a given pathway of genes. The FPE score was based on the enrichment score used in Gene Set Enrichment Analysis (Subramanian et al., 2005). Each gene's coefficients were normalized to obtain the relative importance of each pattern on the gene. Genes which were not common to both the biological process and microarray data after preprocessing were ignored and not included in the computation. Biological processes with fewer than 15 genes were excluded from further analysis. If a gene had multiple probes, the probe which could be reconstructed with least error was chosen. A biological processes' p-value was found by comparing its FPE score to the score of 1000 randomly generated gene sets with the same number of genes. After p-values had been calculated for all processes, they were adjusted for multiple hypotheses testing by using the Benjamini and Hochberg method. A process with a p-value less than 0.05 was considered significant. The leading set of a process is defined as the subset of genes, which was used to compute the Pathway Evaluation score. Genes are not restricted to being in only one leading set allowing for genes to be influenced by multiple patterns and used in multiple functions. The average expression of the genes in the leading set will strongly resemble the original pattern. However, a gene in the leading set is not required to look exactly like the pattern letting allowing for both known biological information and expression patterns to be incorporated.

Leading sets were found from the gene sets of the C2 Canonical Pathways and C5 Gene Ontology databases in MSigDB (Subramanian et al., 2005). The C2 Canonical Pathway database consists of 880 curated sets of genes corresponding to metabolic and signaling pathways. The C5 database consists of 1454 gene sets derived from the Gene Ontology project (http://geneontology.org).

Ingenuity Pathway Analysis. Data were analyzed through the use of Ingenuity Pathway Analysis (IPA) (Ingenuity® Systems, www.ingenuity.com). A network/My Pathway is a graphical representation of the molecular relationships between molecules. Molecules are represented as nodes, and the biological relationship between two nodes is represented as an edge (line). All edges are supported by at least one reference from the literature, from a textbook, or from canonical information stored in the Ingenuity Knowledge Base. Human, mouse, and rat orthologs of a gene are stored as separate objects in the Ingenuity Knowledge Base but are represented as a single node in the network. Nodes are displayed using various shapes that represent the functional class of the gene product.

A total of 773 significant inflammation genes identified in the computational system were subjected to an Inflammatory Response–Inflammation overlay to determine which genes in the significant inflammation leading set were directly involved in inflammation according to IPA (Table 1). A total of 890 significant fibrosis genes were subjected to an Organismal Injury and Abnormalities–Fibrosis overlay to determine which genes in the significant fibrosis leading set were directly involved in fibrosis according to IPA (Table 2). To determine the interactions between genes which have only been experimentally observed in the lung, the Build-Trim tool of IPA was used. Direct and indirect interactions were trimmed to a Confidence Level of Experimentally Observed, and Tissue & Cell Lines included both Organ Systems of Lung and Lung Cell Lines.

Cell culture. Small airway epithelial cells (SAEC) were cultured in SABM media (Lonza) supplemented with a SingleQuot Kit (Lonza). Cells were maintained at 37 °C with 5% CO_2.

Enzyme-linked immunosorbent assay (ELISA). SAEC were plated at 60,000 cells per well in a 24-well dish and grown at 37 °C for 48 h. Cells were serum starved overnight followed by exposure to 1 μg/ml or 2.5 μg/ml MWCNT for 24 h. Conditioned media were collected and assayed for vascular endothelial growth factor A (VEGFA) and C-C motif chemokine 2 (CCL2) protein expression levels using DuoSet ELISA Development Systems from R&D Systems (Minneapolis, MN) according to the manufacturer's protocol. Statistical analysis was done using a two-sample *t*-test assuming unequal variances.

Cellular RNA isolation. RNA was isolated from SAEC using RNAprotect Cell Reagent and an RNeasy Mini Kit from Qiagen according to the manufacturer's protocol (Qiagen, Valencia, CA). RNA concentrations were determined using a NanoDrop 1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE) and RNA quality was

Table 1
Gene information for 34 inflammation genes significantly altered in MWCNT-treated mouse lungs above 1.5 fold change with an FDR of 1% in SAM analysis and with functional molecular interactions in inflammation in the lung in IPA analysis. These genes were also strongly correlated with lung BAL scores in MWCNT-treated mice.

Gene symbol	gene Name	Cellular function	IPA biological functions and disease annotations
ADORA2B	Adenosine A2b receptor	Signal transduction	Proliferation of Cells
C3AR1	Complement component 3a receptor 1	Inflammatory response	Cell Movement, Proliferation of Cells, Morphology of Cells
CCL2	C-C motif chemokine 2	Immune cell chemoattractant	Cell Movement, Proliferation of Cells, Injury of Lung
CCL5	C-C motif chemokine 5	Immune cell chemoattractant	Cell Movement, Proliferation of Cells
CD14	Monocyte differentiation antigen CD14	Innate immune response	Cell Movement, Proliferation of Cells, Injury of Lung
CD86	T-lymphocyte activation antigen CD86	T-cell activation	Cell Movement, Proliferation of Cells,
CD44	CD44 antigen	Cellular adhesion, Immune response	Cell Movement, Proliferation of Cells, Morphology of Cells
CEBPB	CCAAT/enhancer-binding protein beta	Transcription	Proliferation of Cells, Morphology of Cells
CTSS	Cathepsin S	Protease	Cell Movement, Proliferation of Cells
CYBA	Cytochrome b245 light chain	Oxidation	Proliferation of Cells
EGFR	Epidermal growth factor receptor	Signal transduction	Cell Movement, Proliferation of Cells, Morphology of Cells
FCER1G	High affinity immunoglobulin epsilon receptor subunit gamma	Immune response	Cell Movement, Proliferation of Cells, Morphology of Cells
FN1	Fibronectin	Cellular adhesion	Cell Movement,Proliferation of Cells, Degradation of Connective Tissue, Morphology of Cells
IKBKG	NF-kappa-B essential modulator	Kinase activity	Cell Movement, Proliferation of Cells, Morphology of Cells
IL6	Interleukin-6	Immune response	Cell Movement, Proliferation of Cells, Injury of Lung, Degradation of Connective Tissue, Morphology of Cells
IL1B	Interleukin-1 beta	Immune response	Cell Movement, Proliferation of Cells, Degradation of Connective Tissue
IL1R1	Interleukin-1 receptor type 1	Signal transduction	Cell Movement, Proliferation of Cells, Injury of Lung
IL1RN	Interleukin-1 receptor antagonist protein	Interleukin-1 inhibition	Cell Movement, Proliferation of Cells, Degradation of Connective Tissue, Morphology of Cells
IL23A	Interleukin-23 subunit alpha	Immune response	Cell Movement, Proliferation of Cells
JUNB	Transcription factor AP-1	Transcription	Proliferation of Cells, Morphology of Cells
MMP9	Matrix metalloproteinase-9	Peptidase	Cell Movement, Proliferation of Cells, Morphology of Cells
MYD88	Myeloid differentiation primary response protein MYD88		Cell Movement, Proliferation of Cells, Morphology of Cells
NFKBIA	NF-kappa-B inhibitor alpha	Transcription	Cell Movement, Proliferation of Cells, Morphology of Cells
OLR1	Oxidized low-density lipoprotein receptor 1	Signal transduction	Cell Movement, Injury of Lung
PTGS2	Prostoglandin G/H synthase 2	Inflammation	Cell Movement, Proliferation of Cells, Injury of Lung, Degradation of Connective Tissue
SELP	P-selectin	Immune response	Cell Movement, Proliferation of Cells, Injury of Lung, Morphology of Cells
SOCS1	Suppressor of cytokine signaling 1	Signal transduction	Cell Movement, Proliferation of Cells
SOD2	Superoxide dismutase (Mn) mitochondrial	Free radical scavenging	Cell Movement, Proliferation of Cells, Morphology of Cells
SPHK1	Sphingosine kinase 1	Kinase activity	Cell Movement, Proliferation of Cells, Injury of Lung
SPP1	Bone sialoprotein 2	Matrix adhesion	Cell Movement, Proliferation of Cells, Morphology of Cells
TNF	Tumor necrosis factor	Immune response	Cell Movement, Proliferation of Cells, Injury of Lung, Degradation of Connective Tissue, Morphology of Cells
TNFAIP3	Tumor necrosis factor alpha-induced protein 3	Inflammation	Cell Movement, Proliferation of Cells
TNFSF10	Tumor necrosis factor ligand superfamily member 10	Signal transduction	Cell Movement, Proliferation of Cells, Morphology of Cells
VCAM1	Vascular cell adhesion molecule 1	Cell–cell adhesion	Cell Movement, Proliferation of Cells

assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA).

Real-time polymerase chain reaction. Total RNA (1 μg) was converted into complementary DNA (cDNA) using a High Capacity cDNA Reverse Transcription Kit from Applied Biosystems (Life Technologies, Carlsbad, CA). All quantitative real-time PCR (qRT-PCR) reactions were performed on a 7500 Real-Time PCR system from Applied Biosystems. Each treatment group consisted of three biological replicates. qRT-PCR analysis for each biological replicate was performed in triplicate, and the Ct values obtained were normalized to the 18S housekeeping gene. Validated gene expression assays from Applied Biosystems were employed to carry out the mRNA expression profiling. The following gene expression assays were used: VEGFA (Hs00900055_m1); CCL2 (Hs00234140_m1); and 18S (Hs99999901_s1). Thermal cycling conditions were as follows: 50 °C for 2 min, 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 10 min.

Results

Overview of an in vivo MWCNT exposure study and in vitro validation

A schematic of the overall method of determining and validating relevant processes related to lung inflammation and the progression to fibrosis after MWCNT exposure is depicted in Fig. 1A. A total of 480 mice were randomized into three groups 1) gene expression profiling from snap frozen lung tissues, 2) BAL collection for inflammation assessment, and 3) lung tissue fixation for pathological analysis of fibrosis as determined by morphometric analysis of Sirius Red staining of lung tissue for collagen at 1, 7, 28, and 56 days post-exposure to MWCNT by pharyngeal aspiration (Mercer et al., 2011; Porter et al., 2010). Each time point consisted of 8 mice exposed to 0 (dispersion media [DM] control), 10, 20, 40, or 80 μg of MWCNT dispersed in DM for each animal group (Porter et al., 2008, 2010). Genome-wide mRNA expression profiles were analyzed by microarray through mRNA samples purified from the collected tissue and run on an Agilent Mouse Whole Genome Array. BAL was evaluated for the presence of polymorphonuclear leukocytes to assess inflammation, and morphometric analysis of Sirius Red staining for collagen in the alveolar walls was used to evaluate fibrosis (Mercer et al., 2011; Porter et al., 2010). To determine significant transcription-related biological processes and genes with expression corresponding to MWCNT-induced lung inflammation or fibrosis patterns, the computational system (Fig. 1B) evaluated 41,059 probes on the microarray and established biological processes by incorporating the pathological data as input patterns in the simulation, as described in the Materials and methods. For each significant biological process (BH adjusted $p < 0.05$), the corresponding leading set consists of the genes from

Table 2
Gene information for 24 fibrosis genes significantly altered in MWCNT-treated mouse lungs above 1.5 fold change with an FDR of 1% in SAM analysis and with functional molecular interactions in fibrosis in the lung in IPA analysis. These genes were also strongly correlated with morphometric alveolar interstitial fibrosis data in MWCNT-treated mice.

Gene symbol	Gene name	Cellular function	IPA biological functions and disease annotations
ADORA1	Adenosine receptor A1	Signal transduction	Cell Movement, Proliferation of Cells
ADORA2B	Adenosine A2b receptor	Signal transduction	Proliferation of Cells
C3	Complement component 3	Complement system	Cell Movement, Proliferation of Cells, Injury of Lung, Morphology of Cells
CCL2	C-C motif chemokine 2	Immune cell chemoattractant	Cell Movement, Proliferation of Cells, Injury of Lung
CCL17	C-C motif chemokine 17	Immune cell chemoattractant	Cell Movement
CEBPB	CCAAT/enhancer-binding protein beta	Transcription	Proliferation of Cells, Morphology of Cells
FAS	Tumor necrosis factor ligand superfamily member 6	Signal transduction	Cell Movement, Proliferation of Cells, Morphology of Cells
GSK3B	Glycogen synthase kinase-3 beta	Kinase activity	Cell Movement, Proliferation of Cells, Morphology of Cells
IL5	Interleukin-5	Immune response	Cell Movement, Proliferation of Cells, Injury of Lung, Morphology of Cells
IL6	Interleukin-6	Immune response	Cell Movement, Proliferation of Cells, Injury of Lung, Degradation of Connective Tissue, Morphology of Cells
IL11	Interleukin-11	Immune response	Cell Movement, Proliferation of Cells
IL1B	Interleukin-1 beta	Immune response	Cell Movement, Proliferation of Cells, Degradation of Connective Tissue, Morphology of Cells
IL1R1	Interleukin-1 receptor type 1	Signal transduction	Cell Movement, Proliferation of Cells, Injury of Lung
IL1RN	Interleukin-1 receptor antagonist protein	Interleukin-1 inhibition	Cell Movement, Proliferation of Cells, Degradation of Connective Tissue, Morphology of Cells
IRF7	Interferon regulatory factor 7	Transcription	Cell Movement
MMP12	Macrophage metalloelastase	Peptidase	Cell Movement, Proliferation of Cells, Injury of Lung
MYD88	Myeloid differentiation primary response protein MyD88	Immune response	Cell Movement, Proliferation of Cells, Morphology of Cells
PTGS2	Prostoglandin G/H synthase 2	Inflammation	Cell Movement, Proliferation of Cells, Injury of Lung, Degradation of Connective Tissue
SELE	E-selectin	Immune response	Cell Movement, Proliferation of Cells, Morphology of Cells
SELP	P-selectin	Immune response	Cell Movement, Proliferation of Cells, Injury of Lung, Morphology of Cells
SOCS1	Suppressor of cytokine signaling 1	Signal transduction	Cell Movement, Proliferation of Cells
TIMP1	Metalloproteinase inhibitor 1	Proteinase inactivation	Cell Movement, Proliferation of Cells
TNF	Tumor necrosis factor	Immune response	Cell Movement, Proliferation of Cells, Injury of Lung, Degradation of Connective Tissue, Morphology of Cells
TNFAIP3	Tumor necrosis factor alpha-induced protein 3	Inflammation	Cell Movement, Proliferation of Cells

the process which are most strongly related to the input pattern. Not all genes in the leading set will exactly resemble the pattern, but the average expression of the leading set will. The identified significant genes (SAM analysis; $p < 0.05$; FDR <1%; fold change >1.5) in the leading sets were then entered into IPA to identify genes functionally associated with inflammation and fibrosis and to depict molecular interactions in the lung. Based on the comprehensive evaluation, *vegfa* and *ccl2* were selected for *in vitro* validation.

Identification of biological processes with expression patterns resembling MWCNT-induced inflammation or fibrosis pathology

A computational system was used to identify genes and biological processes with transcriptional activities, which matched the observed pathological patterns of lung inflammation or fibrotic collagen in the alveolar wall in the MWCNT-exposed mice (Fig. 1B). The preprocessing step found 2996 unique probes which were significantly ($p < 0.05$; FDR <1%; fold change >1.5) up-regulated (Fig. 2A) or down-regulated (Fig. 2B) using Significance Analysis of Microarrays (SAM) or a linear model showing significant ($p < 0.05$; FDR <0.1%) dose–response or dose and time interactions (Fig. 2C). Using this set of 2996 genes, quantitative BAL and pathological data of MWCNT-induced inflammation or quantitative morphometric analysis of fibrosis were used as input patterns to find gene coefficients for reconstruction of the gene expression. Specifically, results for 3 sets of data were found, 2 sets relating to fibrosis (morphometrically determined changes in collagen within the alveolar wall) (Mercer et al., 2011) and 1 relating to inflammation (BAL) (Porter et al., 2010). Pathology data for fibrosis at dose 80 μg across the 4 time points was fit as an input pattern. The computational system found 64 total significant (BH adjusted $p < 0.05$) leading sets, the subset of genes which was used to compute the Functional Process Evaluation (FPE) score, representing the level of correlation with the fibrosis morphometric data for each biological process in the databases. Morphometric data for fibrosis occurring on day 56 across 4 doses was fit in the computational system with 84 significant (BH adjusted $p < 0.05$) leading sets found. Lastly, inflammation BAL scores at dose 40 μg across 4 time points were used, and 110 leading sets were found to be significantly (BH adjusted $p < 0.05$) correlated with the inflammation pattern.

Example results for each of the pathology data are shown in Fig. 3. The average of the mRNA expression of genes in the leading set closely resembled the pathology data, indicating that in general, the transcriptional activities of the leading set genes correlated with changes in the pathology. The leading sets Reactome GPCR Ligand Binding (Fig. 3A) and Reactome Hemostasis (Fig. 3B) were found in the C2 Canonical Pathway database and consisted of 156 genes (Supplementary data Table 1) and 147 genes (Supplementary data Table 2), respectively. The leading set of Immune System Process (Fig. 3C) was found in the C5 database and consisted of 163 genes (Supplementary data Table 3). CCL2 (Fig. 3D) was contained in the leading set of Reactome GPCR Ligand Binding. Although the CCL2 expression does not exactly follow the pattern, the average of all gene expression in the leading set does. The same can be seen for VEGFA (Fig. 3E). Importantly, our computational system does not constrain genes to being in only one leading set, allowing for genes to be involved in multiple processes. For instance, CCL2 was found to be involved in both MWNCT-induced fibrosis (Fig. 3A) and inflammation (Fig. 3F).

Determination of genes functionally involved in inflammation and fibrosis

To determine which genes were significantly altered in response to MWCNT exposure, leading set genes which attained a fold change of 1.5-fold or greater were input into Ingenuity Pathway Analysis (IPA) to determine if they were functionally involved in inflammation or fibrosis according to currently accepted literature.

The inflammation and fibrosis biological processes consisted of 773 and 890 unique genes, respectively, identified to be significantly altered ($p < 0.05$; fold change >1.5) after MWCNT exposure (significant inflammation) with a false discovery rate (FDR) of 1% in SAM analysis.

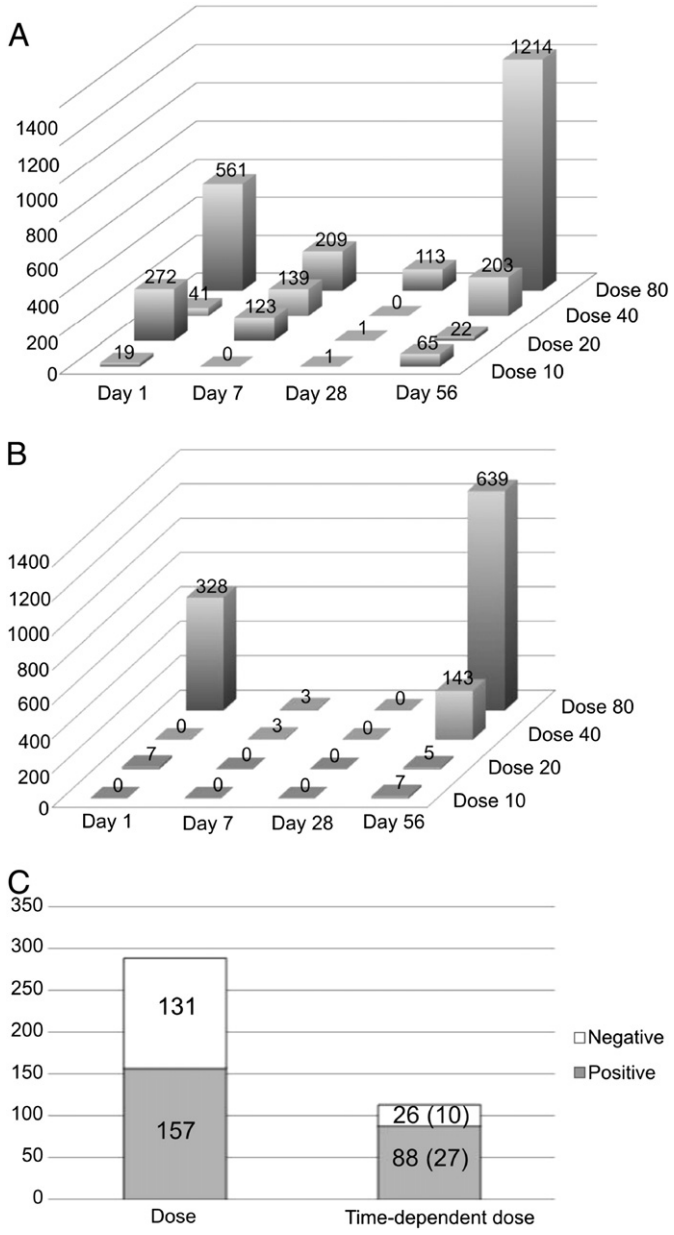

Fig. 2. (A) Number of genes upregulated by using pairwise SAM at each condition. (B) Number of genes down-regulated by using pairwise SAM at each condition. (C) Number of genes significant in the linear model for the dose and dose × time parameters. Numbers in parentheses show the number of genes common to both lists. In all cases for these genes, the parameters were of opposite signs.

Of the 773 significant inflammation genes, 67 were determined to be directly involved in inflammation by IPA (Supplementary data Table 4). Of the 890 significant fibrosis genes, 69 were determined to be directly involved in fibrosis by IPA (Supplementary data Table 5).

A heat map of gene expression for the 67 significant inflammation genes (Fig. 4A) suggested the up- and down-regulation of multiple genes in response to MWCNT exposure. For each gene, the expression fold change and statistical significance at each dose/time condition could be visualized at (http://www.mwcnttranscriptome.org). Overall, expression of *c3ar1*, *fcgr2b*, *pbk*, *pla2g10*, *il2ra*, *il1rn*, *ptgs1*, *cd14*, *igf1*, *ccl2*, *ccl4*, *il1b*, *pla2g7*, *tnfrsf4*, *ghrl*, *slc11a1*, *tnfaip3*, *cd44*, *adora2b*, *gja1*, *tnf*, *ptgs2*, *junb*, *cd86*, *cyba*, *fcer1g*, *ripk3*, and *socs1* was up-regulated on all days at almost all doses. Expression of *itgb2*, *icos*, *il12b*, *ctss*, *ctsd*, *cd48*, and *il21r* was down-regulated at day 1 but increased in expression at almost all doses on days 7 and 28 and all doses on day 56. Expression of *fn1*, *osm*, *selp*, *thbs1*, *pgf*, *tnfsf9*, *adora3*, *il23a*, *myd88*, *il1r1*, *sod2*, *cebpb*, and *nfkbia* was up-regulated at all doses on day 1 with a decrease in expression over time and down-regulation at most doses on day 56. *Spp1* was highly up-regulated on all days, particularly at doses 40 and 80 μg, while *il6* was highly up-regulated on day 1 and had a sustained increase in expression over time. Expression of *ptger3*, *ikbkg*, *cxcl12*, *ccl5*, *tnfsf10*, *card11*, *il24*, *mc2r*, *cort*, *mmp9*, *vcam1*, *agt*, *sphk1*, *app*, *egfr*, and *abcc1* was down-regulated across all days at most doses.

Of the 69 significant fibrosis genes (Fig. 4B), *il1rn*, *lgals3*, *pla2g10*, *ccl17*, *adra2a*, *cxcl12*, *fcgr2b*, *s100a4*, *igf1*, *mx1*, *ccl8*, *arg1*, *mmp13*, *il1b*, *sele*, *hpx*, *timp1*, *ccl2*, *adora2b*, *hmgcr*, *hmgcs1*, *tnfaip3*, *tnfrsf1b*, *adora3*, *c3*, *tnf*, *tpgs2*, and *hif1a* were up-regulated on all days at almost all doses. Expression of *il12b*, *flt3*, *mdk*, *adora1*, and *il2ra* was decreased on day 1 but increased over time, while expression of *pdpn*, *myd88*, *il1r1*, *cebpb*, *mmp14*, *fn1*, *socs1*, *irf7*, *selp*, *osm*, *thbs1*, *oas2*, *ptgir*, and *sstr4* was increased on day 1 and decreased over time. *Il6*, *cxcl10*, *ccr1*, and *mmp12* were highly expressed on day 1 and remained up-regulated over time, while *fas*, *smad4*, *vegfa*, *eif2c1*, *epha2*, *ptk2*, *gsk3b*, *proc*, *f11*, *lyve1*, *pde3a*, *ednrb*, *bdkrb2*, *actc1*, *bmpr2*, and *smurf2* were down-regulated across all days at almost all doses.

Using IPA and these 67 inflammation genes and 69 fibrosis genes, we determined those genes which were significantly involved in IPA Function and Disease Annotations associated with MWCNT-induced fibrosis. A recent report by Mishra et al. (2012) determined that low, physiologically relevant doses of MWCNT equivalent to those in our mouse study could significantly elevate the levels of transforming growth factor β (TGF-β) and matrix metalloproteinase-9 (MMP-9) in lung epithelial cells, as well as increase mechanisms of collagen production and cellular activation. Therefore, we used IPA to determine which genes in our significant (SAM analysis; $p < 0.05$; FDR < 1%; fold change >1.5) inflammation and fibrosis gene sets were involved in these processes (Supplementary data Tables 4 and 5). Many inflammation genes were involved in general cell activation by functional association with the IPA function and disease annotations, including Cell Movement, Proliferation of Cells, and Morphology of Cells (Supplementary data Table 4). Genes found in the significant inflammation set were also involved in the function and disease annotations, including Injury of Lung (*ccl2*, *cd14*, *il6*, *il1r1*, *olr1*, *ptgs1*, *ptgs2*, *selp*, *sphk1*, and *tnf*), Degradation of Connective Tissue (*fn1*, *il6*, *il1b*, *il1rn*, *osm*, *ptgs1*, *ptgs2*, and *tnf*), as well as the signaling pathway VEGF Signaling (*pgf*) (Supplementary data Table 4). No significant inflammatory genes were found in the TGF-β signaling pathway according to IPA. Many fibrosis genes were also involved in the general cell activation function and disease annotations, such as Cell Movement, Proliferation of cells, and Morphology of Cells (Supplementary data Table 5). Several genes in the significant fibrosis set were involved in the function and disease annotations, including Injury of Lung (*adra2a*, *c3*, *ccl2*, *hif1a*, *il5*, *il6*, *il1r1*, *mmp12*, *ptgs2*, *selp*, *tnf*, and *vegfa*), Degradation of Connective Tissue (*fcgr2b*, *fn1*, *il6*, *il1b*, *il1rn*, *mmp13*, *osm*, *ptgir*, *tnf*, and *tnfrsf1b*), as well as the signaling pathway VEGF Signaling (*actc1*, *hif1a*, *ptk2*, and *vegfa*) (Supplementary data Table 5). Interestingly, 3 genes in the significant fibrosis set, *bmpr2*, *smad4*, and *smurf2*, were involved in the IPA TGF-β signaling pathway, again suggesting that TGF-β signaling may play an important role in the progression of fibrosis and that the computational system was efficient in determining those biological processes which were functionally related to MWCNT-induced inflammation and fibrosis. An additional analysis of the significant inflammation (Fig. 5A) and fibrosis (Fig. 5B) genes by IPA determined those genes that have been experimentally shown to have an interaction specifically in the lung (Tables 1 and 2).

VEGFA and CCL2 in vivo and in vitro RNA expression

The inflammation leading set genes (Fig. 5C) and fibrosis leading set genes (Fig. 5D) were ranked by their frequency of inclusion in

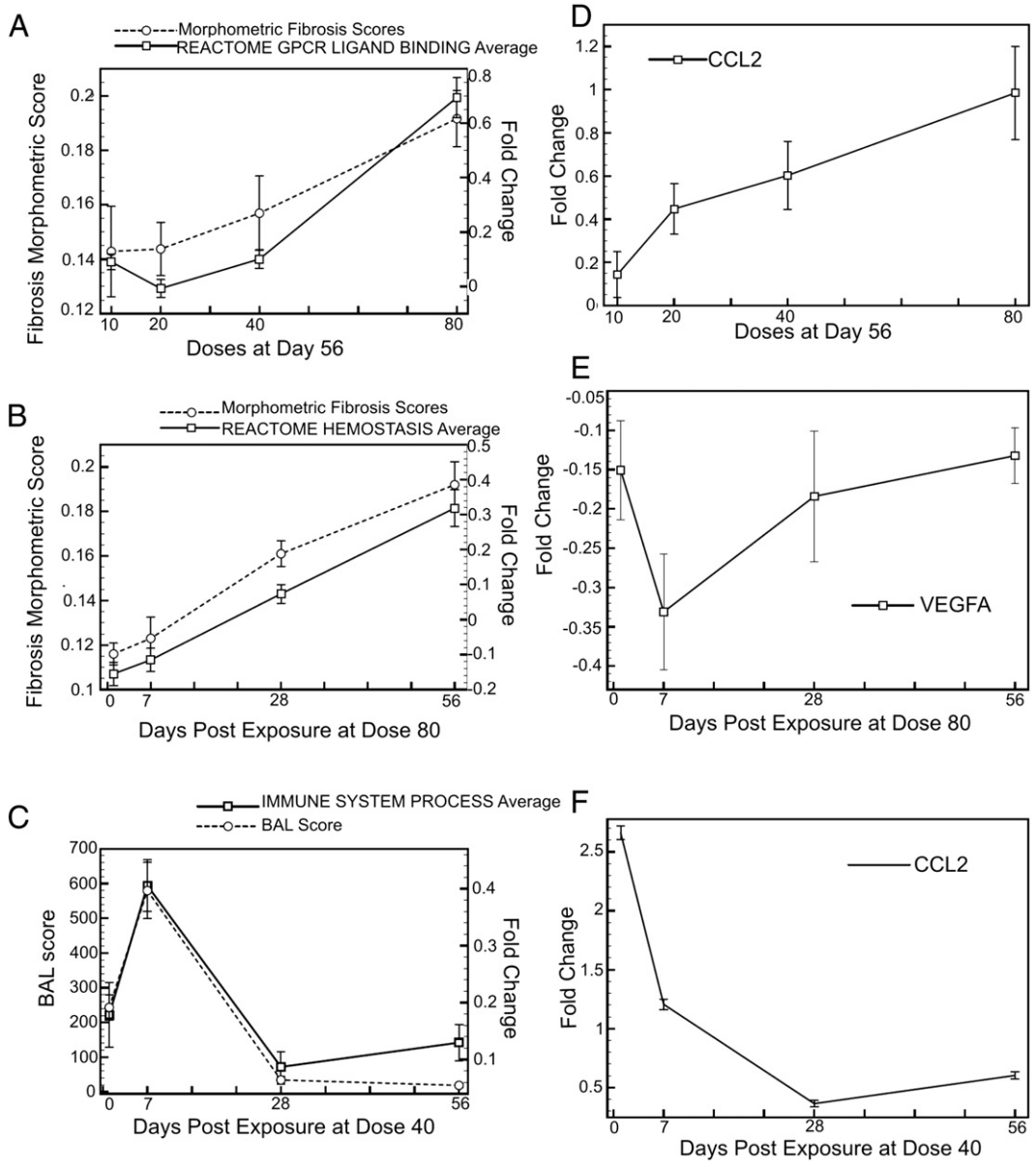

Fig. 3. Three leading sets found to be significant in a search of the C5 and C2 Canonical Pathway databases using pathological data. Computations were based on the observed experimental data points only; lines have been added to emphasize the patterns used in the computational system. For each pathway, (A) Reactome GPCR Ligand Binding, (B) Reactome Hemostasis, and (C) Immune system Process, the average of all the genes in the leading set shows strong similarity to the pathology data. The genes for these leading sets are listed in Supplementary data Tables 1–3. Expression fold change values are shown for CCL2, which was found in the leading sets in (A) and (C), at (D) day 56 and (F) dose 40. VEGFA, found in the leading set from (B), fold change is shown in (E). Although the fold change of these two genes does not exactly match the pathology they were found in the leading sets.

the biological processes significantly correlated with the pathological data. Two genes, *ccl2* and *vegfa*, were selected for *in vitro* validation. *Ccl2* was the top ranked gene that was involved in most biological processes correlated with the inflammation (Fig. 5C) and among the top 20 genes involved in most biological processes correlated with the fibrosis (Fig. 5D). Consistently, in the IPA lung interaction networks (Figs. 5A and B), *ccl2* is in a hub that interacts with both TNF and IL1β hubs in the inflammation and fibrosis networks. *Vegfa* was found to be functionally associated with the fibrosis leading set and is integral to angiogenesis, or the formation of new blood vessels (Ferrara and Davis-Smyth, 1997). Angiogenesis is necessary for the formation of fibrotic tissue, and VEGF has been suggested as a serum biomarker for ranking the severity of idiopathic pulmonary fibrosis (Ando et al., 2010; Strieter and Mehrad, 2009; Thannickal et al., 2004). In a separate study, angiogenesis was observed after MWCNT exposure in human endothelial cells and in a coculture of both human epithelial and endothelial following epithelial exposure (Snyder-Talkington, *In Review*). Both genes were functionally validated in IPA analysis as involved in inflammation and/or fibrosis. Based on these results, *ccl2* and *vegfa* were analyzed for their *in vitro* mRNA and protein expression levels following MWCNT exposure to validate the *in vivo* results. The top ranked gene for fibrosis [chemokine (C motif) receptor 1 (XCR1)] (Fig. 5D) was not shown to be functionally involved in fibrosis in the IPA analysis. Therefore, it was not selected for *in vitro* validation.

In vivo mRNA levels of *vegfa* showed stable expression levels across all days and doses with a significant decrease in expression on day 56 at dose 40 μg (Figs. 5B and 6) and closely resembled the time-course of the morphometric collagen score data and leading set average of the biological process Reactome Hemostasis (Fig. 3B).

Ccl2 showed a consistent dose-dependent increase in mRNA expression on all days with significant increases at all doses on day 1, doses 20, 40, and 80 μg on day 7 and doses 40 and 80 μg on day 56 (Figs. 5A and 6). *Ccl2 in vivo* mRNA expression data closely resembled the fibrosis day 56 dose–response morphometric analysis and leading set average of biological process Reactome GPCR Ligand Binding (Fig. 3A) and was similar to the inflammation BAL pattern and leading set average for Immune System Process (Fig. 3C).

To assess the ability of MWCNT to induce similar RNA expression changes *in vitro*, SAEC were exposed to MWCNT at either 1 μg/ml (approximately equivalent to the *in vivo* dose of 20–40 μg (Porter et al., 2010)) or 2.5 μg/ml (approximately equivalent to the *in vivo* dose of 80 μg (Porter et al., 2010)) for 24 h, and their mRNA expression levels analyzed. MWCNT exposure at both 1 and 2.5 μg/ml exposure levels induced modest but significant increases in *vegfa* mRNA expression *in vitro* in a dose-dependent manner (Fig. 7A). MWCNT exposure at both 1 μg/ml and 2.5 μg/ml levels induced an increase in *ccl2* mRNA expression with a significant increase at 1 μg/ml (Fig. 7B).

VEGFA and CCL2 in vitro protein expression

To determine if the change in *in vitro* mRNA expression levels after exposure to MWCNT resulted in an increase in protein expression, conditioned media from cells exposed to either 1 or 2.5 μg/ml MWCNT for 24 h was collected and analyzed by ELISA for VEGFA and CCL2 protein expression. VEGFA showed significant increases in protein expression levels over control after 24 h of MWCNT exposure (Fig. 7C). CCL2 also showed significant increases in protein expression levels after 24 h of exposure (Fig. 7D). This demonstrated that the increase in mRNA expression levels of VEGFA and CCL2 after MWCNT exposure *in vitro* resulted in a concordant increase in protein expression and indicated that a similar increase may occur after *in vivo* exposure.

Discussion

Integrating *in vivo* and *in vitro* studies and *in silico* analysis is a recent endeavor in toxicological sciences. Novel methods for the analysis of current *in vivo* data are needed to develop predictive *in vitro* models so as to determine the toxicity profile of multiple material variants, such as various types of CNT. Our computational system was sufficient to identify potentially activated functions and pathways, which match inflammatory BAL scores and morphometric alveolar interstitial fibrosis data. By identifying the leading gene sets of the significant functions and pathways, our system can extract genes which are strongly associated with BAL markers and morphometry and that have potential involvement in inflammation and collagen production. The employment of IPA allowed for global analysis of our leading sets throughout the body of accepted scientific literature so as to target our results to those genes known to be

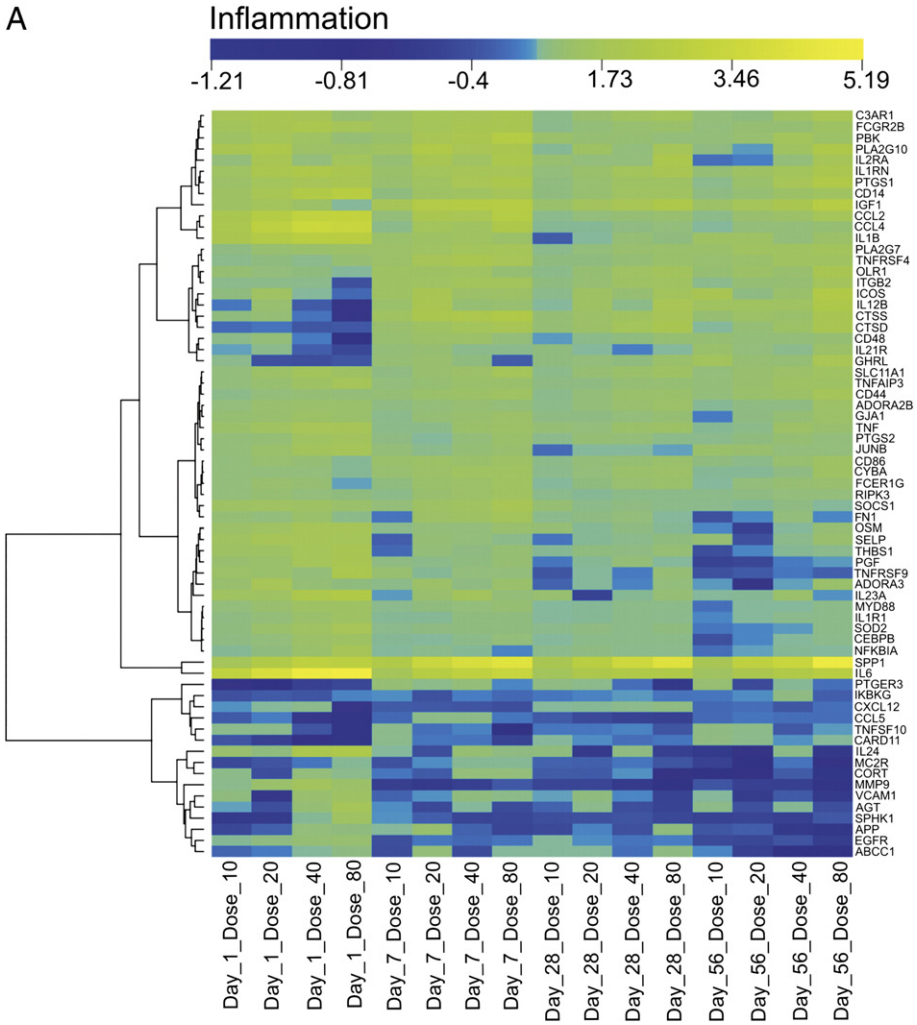

Fig. 4. Heatmap representation of genes significantly altered above 1.5-fold with an FDR of 1% in SAM analysis in inflammation and fibrosis. (A) *In vivo* gene expression of 67 significant inflammation genes across days 1, 7, 28 and 56 at doses 10, 20, 40, and 80 μg. (B) *In vivo* gene expression of 69 significant fibrosis genes across days 1, 7, 28, and 56 at doses 10, 20, 40, and 80 μg.

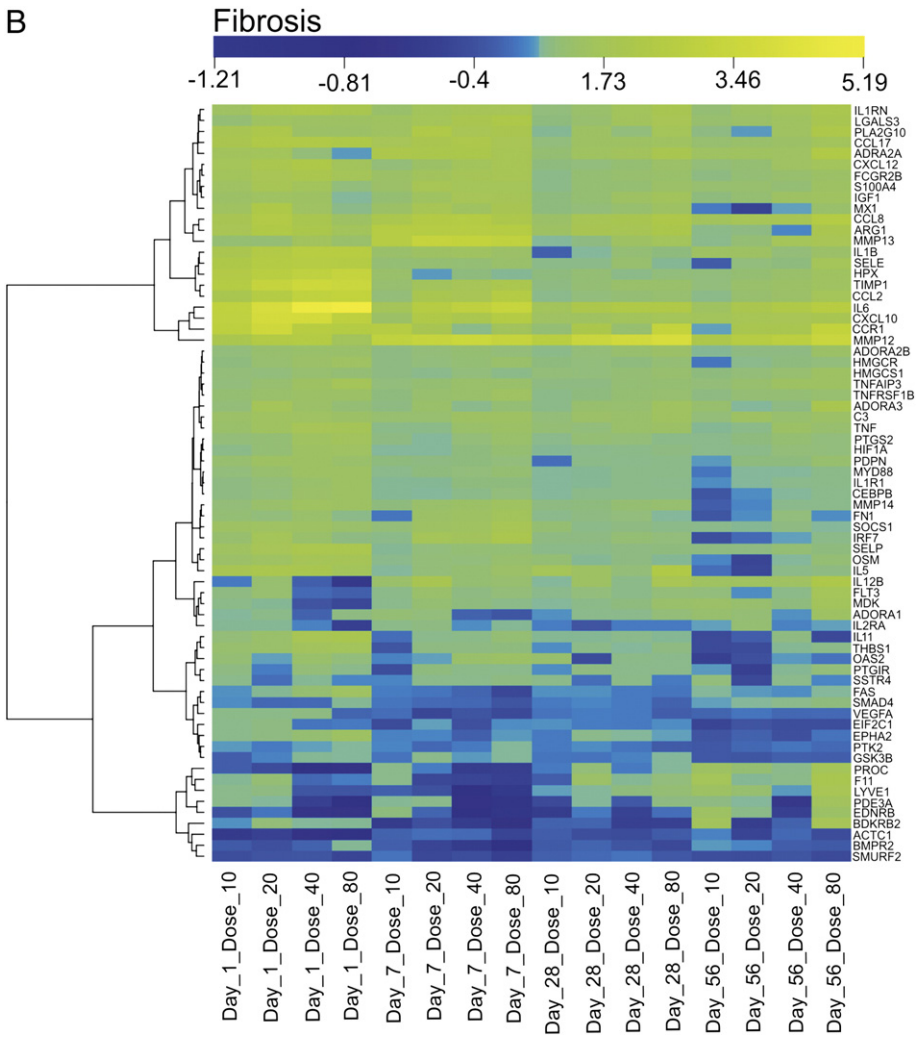

Fig. 4 (*continued*).

involved in inflammation and fibrosis. The comparable results between gene expression profiles of our targeted genes *in vivo* with those found after MWCNT exposure *in vitro* suggest that our computational system is sufficient to determine potential outcomes of MWCNT exposure. This analysis may therefore identify potential signaling pathways and mechanisms that may be studied *in vitro* to determine potential *in vivo* outcomes and prognostic indicators of MWCNT exposure.

The epithelial lining of the lung is the first physical barrier to inhaled particles, and inflammation is a necessary process for the response to and recovery from lung injury. Upon injury, inflammatory mediators are released to trigger an immune response so as to remove the invading pathogen and allow for wound healing, involving processes such as re-epithelialization and angiogenesis (Thannickal et al., 2004; Wynn, 2008). To replace the injured alveolar epithelial cells, alveolar type II cells dedifferentiate and move into the wounded area where they differentiate into new alveolar type I cells. Fibroblast-like cells in the lung secrete extracellular matrix (ECM) along which the alveolar type II cells move (Shi et al., 2009). Upon resolution of the injury, this ECM is typically reabsorbed and normal lung function and architecture are maintained; however, when an inflammatory response becomes chronic due to persistent injury or uncontrolled signaling, the inflammatory process can become pathogenic. Fibrosis, excessive collagen production, in the lung occurs when the deposition of ECM is poorly controlled and there is a loss of normal lung function and architecture (Shi et al., 2009). Because many of the regulators of normal development and inflammation also govern the process of fibrosis, there are multiple hypotheses on what actions must occur for the response to switch from normal inflammation to pathogenic fibrosis (Strieter and Mehrad, 2009).

The exposure of mice by pharyngeal aspiration to 0, 10, 20, 40, and 80 μg of MWCNT was suggested to represent an exposure that could be compared to human occupational exposures (Porter et al., 2010). Sampling of lungs at 1, 7, 28, and 56 days post-exposure allowed the determination of gene expression changes that occur in the initial inflammatory stage as well as in later fibrotic stages of disease. These studies suggested that a single exposure to MWCNT not only induce inflammation, but that MWCNT are biopersistent and induce a delayed fibrotic response as determined by increased collagen in the alveolar walls. Using a novel computational system, the correlation of global mRNA expression profiles to the changes in BAL score and morphometric analysis was analyzed. This identified transcription-related biological processes with expression patterns resembling the pathological patterns of inflammation and fibrosis in MWCNT-exposed mice, allowing for the identification of critical toxicity pathways and potential mechanisms for intervention. The results showed that this systematic analysis could identify relevant genes and pathways in MWCNT-induced lung injury from *in vivo* studies, which were further validated in *in vitro* experiments.

Previous studies used *in vivo* or *in vitro* genome-wide mRNA expression data to infer toxicity in carbon nanotube-exposed rats (Alazzam et al., 2010; Ellinger-Ziegelbauer and Pauluhn, 2009; Peng

et al., 2010). In addition, a combination of microarray data, benchmark dose methods, and Gene Ontology annotations were used to identify potentially adverse biological processes in toxicity (Burgoon and Zacharewski, 2008; Thomas et al., 2007). Our novel computational system allowed for the discovery of non-parametric patterns, which could be used to reconstruct microarray data, incorporated quantitative pathological data, and was capable of working on both time-series and dose-dependent data. Unlike traditional clustering techniques (Tamayo

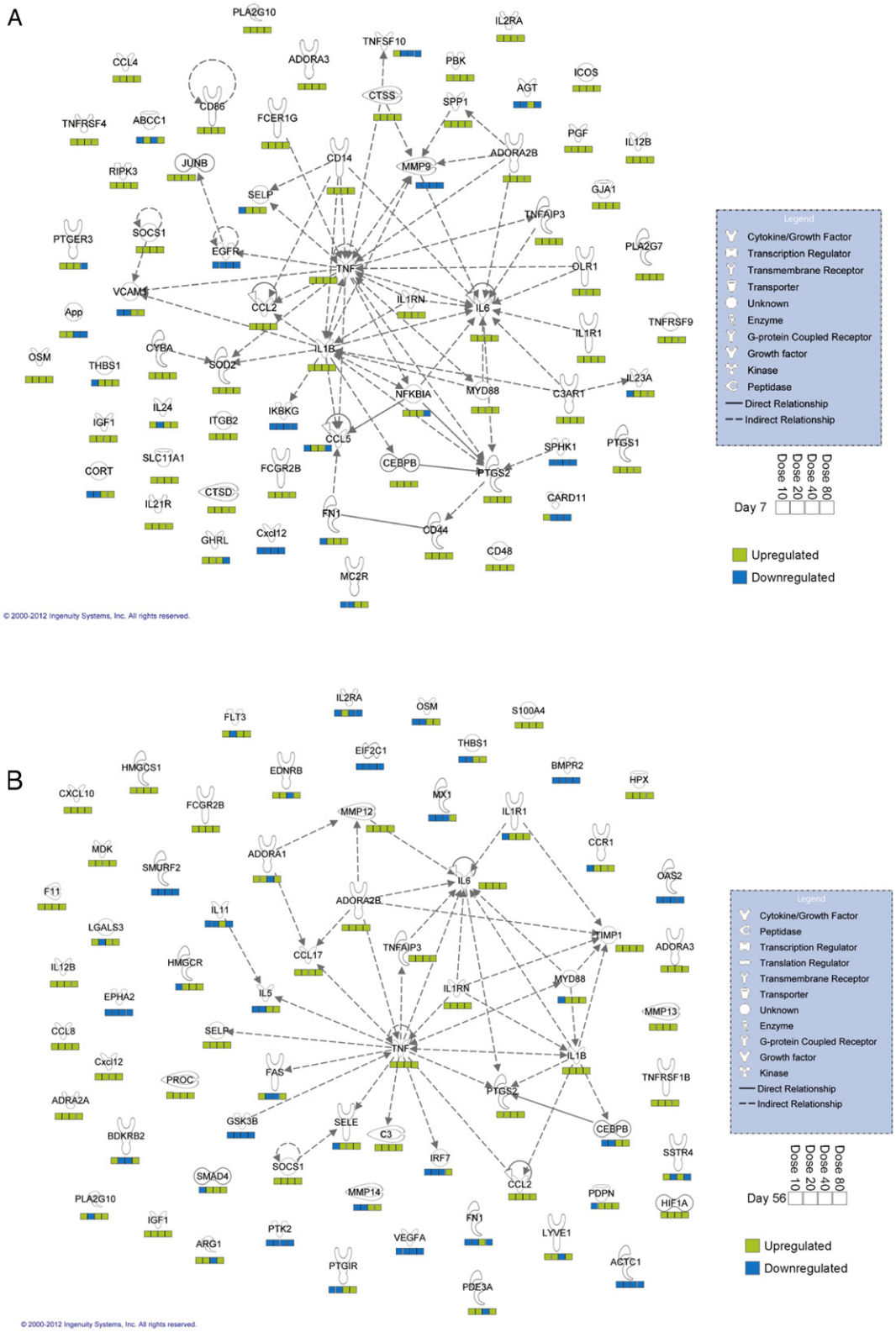

Fig. 5. (A) IPA analysis of the 67 significant inflammation genes to determine those interactions, which specifically occur in the lung. (B) IPA analysis of the 69 significant fibrosis genes to determine those interactions, which specifically occur in the lung. (C) Ranking of significant inflammation genes by their frequency of appearance in biological processes significantly correlated with histopathological data. (D) Ranking of significant fibrosis genes by their frequency of appearance in biological processes significantly correlated with histopathological data.

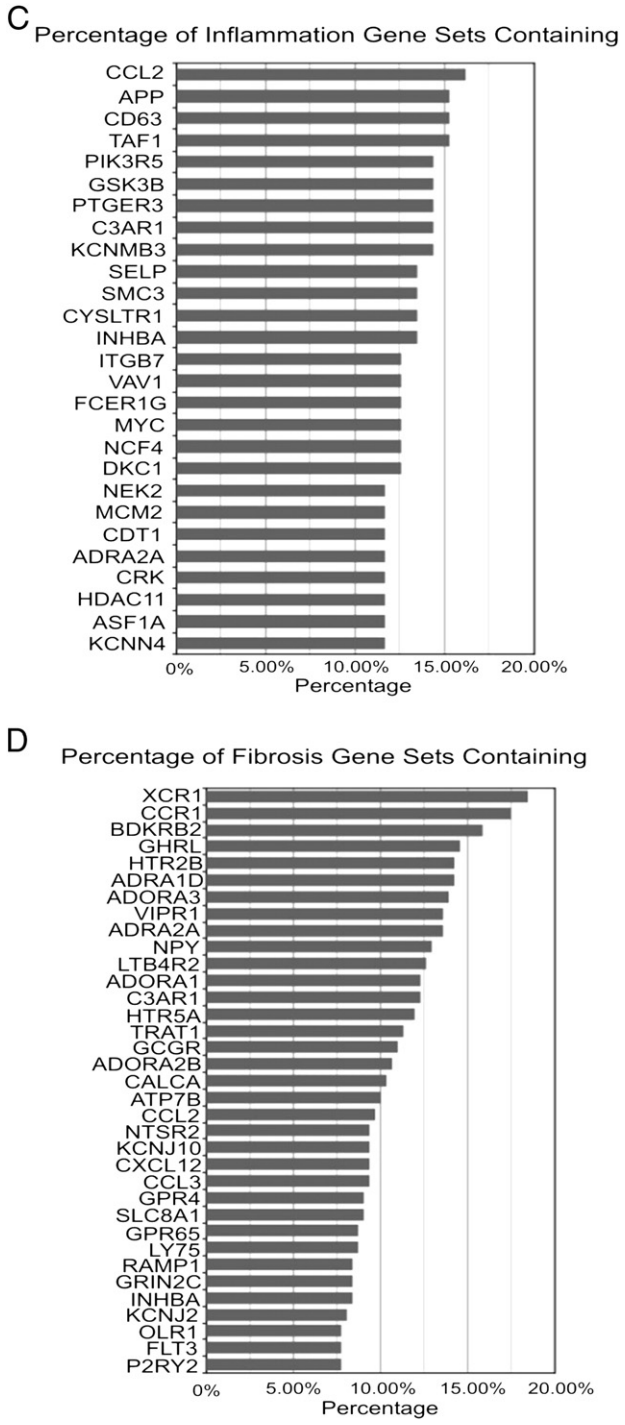

Fig. 5 (*continued*).

et al., 1999; Waring et al., 2001; Yeung and Ruzzo, 2001), our system enabled genes to be included in multiple coexpression groups and be involved with multiple patterns. As a result, 23 genes were found to be involved in both MWCNT-induced inflammation and fibrosis in this study (Tables 1 and 2). The leading sets could be thought of as the genes in a function or pathway, which were most strongly influenced by a particular pattern. Although the expression of an individual gene in the leading set may not exactly match a pattern, the average of all the gene expression in the leading set will, indicating that the process as a whole responds similarly to the pattern. In this study, only genes in the leading sets related to inflammation or fibrosis and from gene sets which were found to be significantly represented were studied. It should be noted that similar to the Bayesian Decomposition method, our computational system uses data collected from multiple conditions, in this study either 4 time points or 4 dose conditions, for each pattern matching and gene expression reconstruction. This computational model would not be applicable to modeling experimental data collected with less than 3 treatment conditions. In addition, since our computational model is non-parametric, it does not make any inference of unobserved experimental conditions.

The use of IPA to determine if genes significantly altered in the leading sets were involved in inflammation or fibrosis allowed for an in depth analysis based upon data derived from relationships between genes and disease states taken from the currently accepted literature knowledge base. These analyses were rooted in and verified by experimental results collated from numerous sources. A total of 67 significantly altered genes were determined by IPA to be directly involved in the inflammatory process while 69 significantly altered genes were determine by IPA to be directly involved in fibrosis. Of the significantly altered genes, two genes, *ccl2* and *vegfa*, were chosen to determine their *in vivo* and *in vitro* expression levels due to their roles in the cell during the development of inflammation and fibrosis as well as their rankings during gene profiling.

The formation of new blood vessels is an early response to tissue injury and a continuous process in the formation of fibrosis. Angiogenesis is necessary to sustain the tissue with oxygen, and increased angiogenic potential has been seen in patients with fibrosis (Strieter, 2008; Thannickal et al., 2004). VEGFA is a predominant angiogenic factor that acts upon endothelial cells for the proliferation of new blood vessels (Ferrara and Davis-Smyth, 1997). Because angiogenesis is integral to the formation of excessive ECM, we chose to determine if MWCNT had the ability to increase the expression of *vegfa* for both inflammatory and fibrotic processes. Additionally, we chose to determine the expression levels of *ccl2* as an indicator of the inflammatory process. CCL2 is a known stimulator of the immune response, initiating chemotaxis in a variety of cell types, such as monocytes, lymphocytes, and basophils, as well as inducing the production of collagen from fibroblast cells. Due to these processes, CCL2 is suggested to play a role in inflammatory diseases (Rose et al., 2003). Interestingly, CCL2 has also been suggested to play a role in angiogenesis and up-regulates the expression of VEGFA while, in turn, VEGFA has also been suggested to increase the expression of CCL2 (Hong et al., 2005; Yadav et al., 2010).

The dose-dependent increase in *ccl2* mRNA expression at all days and doses *in vivo* suggests its role in the initial inflammatory process. Although the *in vivo* mRNA levels of *vegfa* remained relatively constant across all days and doses, the *in vivo* protein levels are unknown and may enhance collagen production. *In vitro* levels of *ccl2* and *vegfa* mRNA also increased with increasing dose, reflecting what is seen in the *in vivo* analysis. *In vitro* analysis of the protein levels of CCL2 and VEGFA suggests that even modest increases in mRNA levels were able to significantly up-regulate protein expression, and a similar increase in protein expression may occur *in vivo*. The analogous changes to *vegfa* mRNA levels *in vitro*, with subsequent increases in protein levels, suggest that MWCNT may have a similar effect *in vitro* to that seen *in vivo*. This may allow for potentially significant cellular processes to be identified by computational means and for the analysis of the mechanisms and signaling cascades behind MWCNT-induced effects to be validated in an *in vitro* manner.

Conclusions

A novel computational model was presented which was sufficient to determine transcription-related biological processes strongly associated with BAL and morphometric markers of lung inflammation and fibrosis, respectively, following exposure to MWCNT in mice. The biological processes were analyzed through IPA to determine genes and signaling pathways functionally involved in lung inflammation and fibrosis. Concordance of expression in two representative functionally

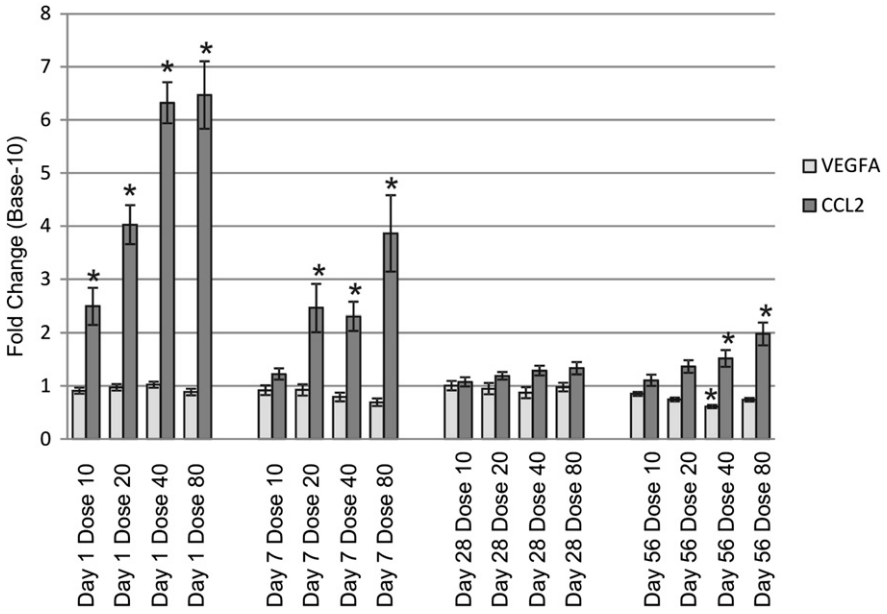

Fig. 6. Base-10 fold change of *in vivo* gene expression data of *vegfa* and *ccl2* on days 1, 7, 28, and 56 at doses 10, 20, 40, and 80 μg. * fold change greater than 1.5.

involved genes in the *in vivo* analysis was confirmed *in vitro*, and the novel computational model was validated as a useful method to identify potential toxicity pathways. The use of these toxicogenomic data and *in vivo* animal model-based gene expression profiling integrated with *in vitro* verification may allow for successful *in vitro* toxicity profiling of MWCNT as well as the identification of potential signaling pathways involved in the etiology of MWCNT-induced injury.

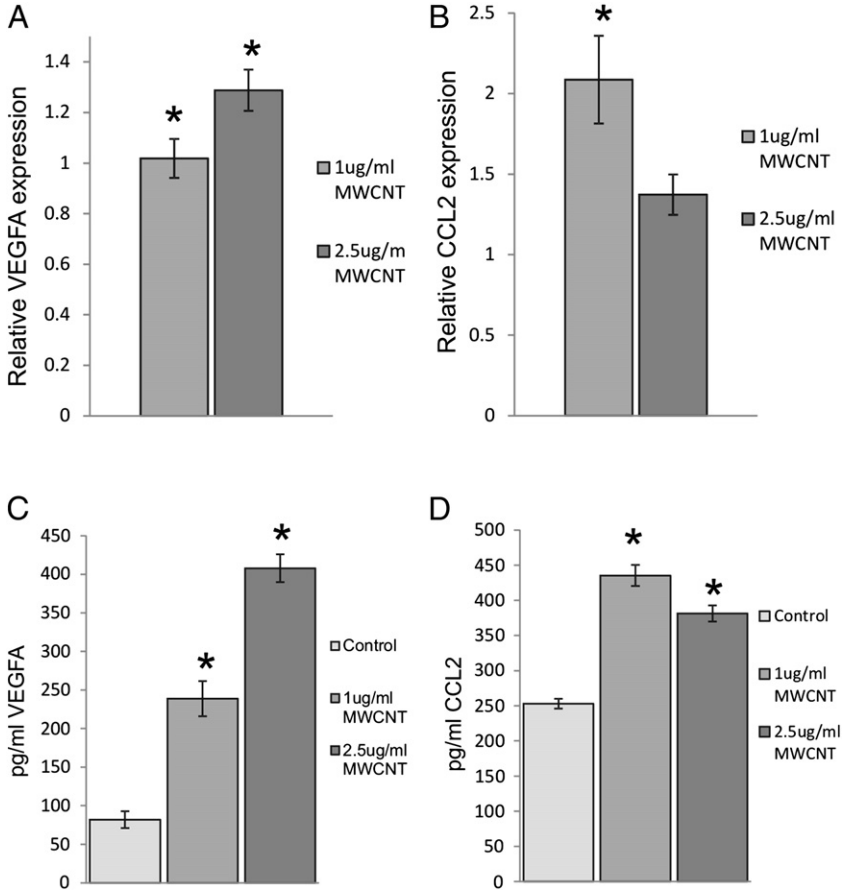

Fig. 7. *In vitro* gene and protein expression data of VEGFA and CCL2. (A) *In vitro vegfa* gene expression in SAEC after exposure to 1 μg/ml and 2.5 μg/ml MWCNT for 24 h. (B). *In vitro ccl2* gene expression in SAEC exposure to 1 μg/ml and 2.5 μg/ml MWCNT for 24 h. (C) ELISA results of VEGFA protein expression after DM (82 ± 11 pg/ml), 1 μg/ml (239 ± 23 pg/ml), and 2.5 μg/ml (408 ± 18 pg/ml) MWCNT exposure for 24 h. (D) ELISA results of CCL2 protein expression after DM (253 ± 7 pg/ml), 1 μg/ml (435 ± 15 pg/ml), and 2.5 μg/ml (381 ± 11 pg/ml) MWCNT exposure for 24 h. * $p < 0.05$.

Financial support

Dr. Nancy L. Guo is supported by NIH R01ES021764; R01/R56LM009500. Julian Dymacek is supported by a NSF training grant on nanotoxicology. Dr. James Denvir was supported by awards from the National Center for Research Resources/National Institute of General Medical Sciences (2P20RR016477/8P20GM103434) to the WV-IDeA Network of Biomedical Research Excellence (WV-INBRE).

Disclaimer

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

List of abbreviations

MWCNT	multi-walled carbon nanotubes
IPA	Ingenuity Pathway Analysis
VEGFA	vascular endothelial growth factor A
CCL2	C-C motif chemokine 2
ROS	reactive oxygen species
TEM	transmission electron microscopy
DM	dispersion media
CE	Coefficient Expander
FPE	functional process evaluation
SAM	significance of analysis of microarrays
SAEC	small airway epithelial cells
ELISA	enzyme-linked immunosorbent assay
qRT-PCR	quantitative real-time polymerase chain reaction
BAL	bronchoalveolar lavage

Conflict of interest statement

The author(s) declare they have no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.taap.2013.06.026.

References

Ajayan, P.M., 1999. Nanotubes from carbon. Chem. Rev. 99, 1787–1800.
Alazzam, A., Mfoumou, E., Stiharu, I., Kassab, A., Darnel, A., Yasmeen, A., Sivakumar, N., Bhat, R., Al Moustafa, A.E., 2010. Identification of deregulated genes by single wall carbon-nanotubes in human normal bronchial epithelial cells. Nanomedicine 6, 563–569.
Ando, M., Miyazaki, E., Ito, T., Hiroshige, S., Nureki, S.I., Ueno, T., Takenaka, R., Fukami, T., Kumamoto, T., 2010. Significance of serum vascular endothelial growth factor level in patients with idiopathic pulmonary fibrosis. Lung 188, 247–252.
Burgoon, L.D., Zacharewski, T.R., 2008. Automated quantitative dose-response modeling and point of departure determination for large toxicogenomic and high-throughput screening data sets. Toxicol. Sci. 104, 412–418.
Castranova, V., 2011. Overview of current toxicological knowledge of engineered nanoparticles. J. Occup. Environ. Med. 53, S14–S17.
Dymacek, J., Guo, N.L., 2011. Systems approach to identifying relevant pathways from phenotype information in dose-dependent time series microarray data. IEEE Interanational Conference on Bioinformatics and Biomedicine, pp. 290–293.
Ellinger-Ziegelbauer, H., Pauluhn, J., 2009. Pulmonary toxicity of multi-walled carbon nanotubes (Baytubes) relative to alpha-quartz following a single 6 h inhalation exposure of rats and a 3 months post-exposure period. Toxicology 266, 16–29.
Ferrara, N., Davis-Smyth, T., 1997. The biology of vascular endothelial growth factor. Endocr. Rev. 18, 4–25.
Guo, N.L., Wan, Y.W., Denvir, J., Porter, D.W., Pacurari, M., Wolfarth, M.G., Castranova, V., Qian, Y., 2012. Multiwalled carbon nanotube-induced gene signatures in the mouse lung: potential predictive value for human lung cancer risk and prognosis. J. Toxicol. Environ. Health A 75, 1129–1153.
He, X., Young, S.H., Schwegler-Berry, D., Chisholm, W.P., Fernback, J.E., Ma, Q., 2011. Multiwalled carbon nanotubes induce a fibrogenic response by stimulating reactive oxygen species production, activating NF-kappaB signaling, and promoting fibroblast-to-myofibroblast transformation. Chem. Res. Toxicol. 24, 2237–2248.
Hong, K.H., Ryu, J., Han, K.H., 2005. Monocyte chemoattractant protein-1-induced angiogenesis is mediated by vascular endothelial growth factor-A. Blood 105, 1405–1407.
Iijima, S., 1991. Helical microtubules of graphitic carbon. Nature 354, 56–58.
Junqueira, L.C., Bignolas, G., Brentani, R.R., 1979. Picrosirius staining plus polarization microscopy, a specific method for collagen detection in tissue sections. Histochem. J. 11, 447–455.
Mercer, R.R., Hubbs, A.F., Scabilloni, J.F., Wang, L., Battelli, L.A., Schwegler-Berry, D., Castranova, V., Porter, D.W., 2010. Distribution and persistence of pleural penetrations by multi-walled carbon nanotubes. Part. Fibre Toxicol. 7, 28.
Mercer, R.R., Hubbs, A.F., Scabilloni, J.F., Wang, L., Battelli, L.A., Friend, S., Castranova, V., Porter, D.W., 2011. Pulmonary fibrotic response to aspiration of multi-walled carbon nanotubes. Part. Fibre Toxicol. 8, 21.
Mishra, A., Rojanasakul, Y., Chen, B.T., Castranova, V., Mercer, R.R., Wang, L.Y., 2012. Assessment of Pulmonary Fibrogenic Potential of Multiwalled Carbon Nanotubes in Human Lung Cells. J. Nanomater. 2012, 18. http://dx.doi.org/10.1155/2012/930931 (Article ID 930931).
Muller, J., Huaux, F., Moreau, N., Misson, P., Heilier, J.F., Delos, M., Arras, M., Fonseca, A., Nagy, J.B., Lison, D., 2005. Respiratory toxicity of multi-wall carbon nanotubes. Toxicol. Appl. Pharmacol. 207, 221–231.
Oberdorster, G., Oberdorster, E., Oberdorster, J., 2005. Nanotoxicology: an emerging discipline evolving from studies of ultrafine particles. Environ. Health Perspect. 113, 823–839.
Pacurari, M., Qian, Y., Porter, D.W., Wolfarth, M., Wan, Y., Luo, D., Ding, M., Castranova, V., Guo, N.L., 2011. Multi-walled carbon nanotube-induced gene expression in the mouse lung: association with lung pathology. Toxicol. Appl. Pharmacol. 255, 18–31.
Pacurari, M., Qian, Y., Fu, W., Schwegler Berry, D., Ding, M., Castranova, V., Guo, N.L., 2012. Cell permeability, migration, and reactive oxygen species induced by multiwalled carbon nanotubes in human microvascular endothelial cells. J. Toxicol. Environ. Health A 75, 112–128.
Peng, L., Barczak, A.J., Barbeau, R.A., Xiao, Y., LaTempa, T.J., Grimes, C.A., Desai, T.A., 2010. Whole genome expression analysis reveals differential effects of TiO2 nanotubes on vascular cells. Nano Lett. 10, 143–148.
Porter, D., Sriram, K., Wolfarth, M., Jefferson, A., Schwegler-Berry, D., Andrew, M., Castranova, V., 2008. A biocompatible medium for nanoparticle dispersion. Nanotoxicology 2, 144–154.
Porter, D.W., Hubbs, A.F., Mercer, R.R., Wu, N., Wolfarth, M.G., Sriram, K., Leonard, S., Battelli, L., Schwegler-Berry, D., Friend, S., Andrew, M., Chen, B.T., Tsuruoka, S., Endo, M., Castranova, V., 2010. Mouse pulmonary dose- and time course-responses induced by exposure to multi-walled carbon nanotubes. Toxicology 269, 136–147.
Rose Jr., C.E., Sung, S.S., Fu, S.M., 2003. Significant involvement of CCL2 (MCP-1) in inflammatory disorders of the lung. Microcirculation 10, 273–288.
Shi, W., Xu, J., Warburton, D., 2009. Development, repair and fibrosis: what is common and why it matters. Respirology 14, 656–665.
Srivastava, R.K., Pant, A.B., Kashyap, M.P., Kumar, V., Lohani, M., Jonas, L., Rahman, Q., 2011. Multi-walled carbon nanotubes induce oxidative stress and apoptosis in human lung cancer cell line-A549. Nanotoxicology 5, 195–207.
Strieter, R.M., 2008. What differentiates normal lung repair and fibrosis? Inflammation, resolution of repair, and fibrosis. Proc. Am. Thorac. Soc. 5, 305–310.
Strieter, R.M., Mehrad, B., 2009. New mechanisms of pulmonary fibrosis. Chest 136, 1364–1370.
Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., Mesirov, J.P., 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A. 102, 15545–15550.
Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E.S., Golub, T.R., 1999. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. Proc. Natl. Acad. Sci. U. S. A. 96, 2907–2912.
Thannickal, V.J., Toews, G.B., White, E.S., Lynch III, J.P., Martinez, F.J., 2004. Mechanisms of pulmonary fibrosis. Annu. Rev. Med. 55, 395–417.
Thomas, R.S., Allen, B.C., Nong, A., Yang, L., Bermudez, E., Clewell III, H.J., Andersen, M.E., 2007. A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure. Toxicol. Sci. 98, 240–248.
Walker, V.G., Li, Z., Hulderman, T., Schwegler-Berry, D., Kashon, M.L., Simeonova, P.P., 2009. Potential *in vitro* effects of carbon nanotubes on human aortic endothelial cells. Toxicol. Appl. Pharmacol. 236, 319–328.
Waring, J.F., Ciurlionis, R., Jolly, R.A., Heindel, M., Ulrich, R.G., 2001. Microarray analysis of hepatotoxins *in vitro* reveals a correlation between gene expression profiles and mechanisms of toxicity. Toxicol. Lett. 120, 359–368.
Weibel, E., 1980a. Stereological Methods: Practical Methods for Biological Morphometry, 1st ed. Academic Press, New York, NY.
Weibel, E., 1980b. Stereological Methods: Theoretical Foundations, 1st ed. Academic Press, New York, NY.
Wynn, T.A., 2008. Cellular and molecular mechanisms of fibrosis. J. Pathol. 214, 199–210.
Yadav, A., Saini, V., Arora, S., 2010. MCP-1: chemoattractant with a role beyond immunity: a review. Clin. Chim. Acta 411, 1570–1579.
Ye, S.F., Wu, Y.H., Hou, Z.Q., Zhang, Q.Q., 2009. ROS and NF-kappaB are involved in upregulation of IL-8 in A549 cells exposed to multi-walled carbon nanotubes. Biochem. Biophys. Res. Commun. 379, 643–648.
Yeung, K.Y., Ruzzo, W.L., 2001. Principal component analysis for clustering gene expression data. Bioinformatics 17, 763–774.