Microarray-based Analysis of Ventilator-induced Lung Injury

Mark M. Wurfel

Division of Pulmonary and Critical Care Medicine, Harborview Medical Center, University of Washington, Seattle, Washington

Acute respiratory distress syndrome (ARDS) and acute lung injury (ALI) are a frequent cause of intensive care unit admission, affecting over 200,000 patients in the United States each year. Mechanical ventilation is a life-saving intervention in the setting of ARDS and ALI, but clinical trials have demonstrated that mechanical ventilation with excessive tidal volumes plays a role in promoting and perpetuating lung injury and leads to excess mortality. This process has been labeled ventilator-induced lung injury (VILI), but the molecular mechanisms driving this process and its interactions with predisposing risk factors such as sepsis and chemical injury remain incompletely understood. Genome-wide measurements of gene expression using microarray technology represent a powerful tool to examine the pathophysiology of VILI. Several recent studies have used this approach to study VILI in isolation and associated with endotoxin instillation or saline lavage. These studies and others examining gene expression profiles in epithelial cells subjected to cyclic stretch have provided novel insights on the molecular mechanisms underlying VILI. This review will summarize these findings and discuss implications for future studies.

Keywords: microarray analysis; lung injury; mechanical ventilation; gene expression; genomics

VENTILATOR-INDUCED LUNG INJURY

As early as three decades ago, it was recognized in animal models that mechanical ventilation with high tidal volumes is injurious to the lung and can cause alveolar disruption and pulmonary edema (1). However, it was not until the ARDS network trial demonstrating that lowering the tidal volumes delivered to mechanically ventilated patients with acute lung injury reduces the degree of lung injury and mortality (2) that it became clear that mechanical ventilation plays an important role in the propagation of lung injury, a process referred to as ventilator-induced lung injury (VILI) (3). Despite this demonstration of the deleterious effects of excessive tidal volumes in ARDS and ALI, there is, as of yet, incomplete understanding of the physiologic and molecular mechanisms underlying VILI. The diversity of predisposing risk factors for ALI, such as sepsis, pneumonia, trauma, and massive transfusion, and the ill-defined molecular definition of ALI, further confounds our ability to elucidate these mechanisms. Recent advances capitalizing on the elucidation of the human and rodent “transcriptome” have provided an important tool with which to define and classify VILI.

OVERVIEW OF MICROARRAY-BASED GENE EXPRESSION ANALYSES

Microarray-based genome-wide expression analyses can provide an unbiased view of the transcriptional state of a cell population (4, 5). DNA microarray analysis was first reported in 1995 as a method to determine the expression of thousands of genes simultaneously (5, 6). Microarray platforms and analysis technology have undergone major advances over the past decade since its inception (reviewed in References 7 and 8). Application of microarray analyses to human disease pathogenesis has resulted in the discovery of new sub-classes of neoplasms that confer differing prognosis or response to therapies (9–12). Microarray-based expression profiling has also uncovered new pathogenic mechanisms involved in pulmonary diseases such as asthma, pulmonary fibrosis, and pulmonary hypertension (13–15).

Gene expression data obtained from the most current generation of microarray platforms, such as photolithograph-based oligonucleotide microarrays, have been shown to produce reliable measurements in experienced laboratories (16). This microarray platform produces expression data that shows minimal technical variation (i.e., differences between measurements made from the same starting RNA sample on different arrays) as measured by expression value correlation and agreement, while showing high correlation ($r = 0.8$ – 0.9) with measurements from “gold standard” approaches such as quantitative real-time PCR (QPCR) (16, 17). These studies show that gene expression analyses using state-of-the-art microarray platforms give valid estimates of the expression levels of most genes.

In contrast to the estimation of true gene expression values, the measurements of gene expression obtained by microarray are only moderately correlated ($r = 0.5$ – 0.6) with expression levels of corresponding protein products (18, 19). This could be due to post-transcriptional regulation such as splice variations, mRNA stability, translation initiation, and protein stability. Thus, extrapolation of protein expression on the basis of microarray-based gene expression measurements must be validated by direct measurement of the protein. Direct proteomic analyses might eventually replace microarray-based gene expression analyses in some experimental systems, although current proteomic approaches remain much more cumbersome and expensive than genomic approaches.

Quantitation of the sources of experimental variability in microarray experiments using oligonucleotide array-based profiling of tissue samples shows that the largest component of variability, or “noise,” is introduced by biological effects in both human (20) and murine (21) samples. The sources for this biological noise could include inter-individual differences in environment (e.g., diet, animal stress), tissue sample handling (e.g., time until RNA purification), inadequate sampling of heterogeneous tissue (e.g., lung, liver), and, in outbred populations such as humans, underlying genetic variation. The main implication of these findings for experimental design of microarray studies involving gene expression profiling of tissues derived from whole animals is that biological replicates are imperative. Furthermore, this implies that pooling of biologic replicates, a technique often used to minimize the amount of costly microarray reagents used

(Received in original form August 1, 2006; accepted in final form September 12, 2006)

This study was supported by NHLBI grant K23 HL72923.

Correspondence and requests for reprints should be addressed to Mark M. Wurfel, Harborview Medical Center, Research and Training Bldg, Rm 606, Box 359640, 300 9th Ave., Seattle, WA, 98104. E-mail: mwurfel@u.washington.edu

Proc Am Thorac Soc Vol 4, pp 77–84, 2007

DOI: 10.1513/pats.200608-149JG

Internet address: www.atsjournals.org

in an experiment, greatly reduces the statistical confidence in gene expression values obtained, given that biological noise is not accounted for.

The number of different statistical tests available for use in the analysis of microarray data has greatly expanded over the past decade. For an overview of state-of-the-art analytic issues and approaches *see* the review by Allison and coworkers (7). In brief, the critical components to the microarray data analysis pipeline are: (1) signal normalization, (2) statistical filtering, (3) correction for multiple hypothesis testing, and (4) data visualization and mining. Correction for multiple hypothesis testing is a particularly important problem in microarray analysis given that thousands of gene expression “hypotheses” are tested on each array. This means that, at standard levels of statistical significance (e.g., 0.05), an array testing expression of 10,000 genes could call 500 genes as differentially regulated purely by chance and would, therefore, be false positives or type I error (rejection of null hypothesis when actually true). Methods to minimize this error range from the overly conservative Bonferroni correction to the more recently developed programs estimating the false discovery rate (FDR), the expected fraction of null hypotheses rejected mistakenly (22). The latter test is robust and, in contrast to most other multiple comparisons corrections, can be applied to correlated data structures (22) such as exist in most microarray experiments. This approach has become widely accepted in the microarray literature, and all of the articles described below use some form of this approach in their analyses. Overall, genome-wide measurements of gene expression have become an important tool in the dissection of pathophysiologic mechanisms and discovery of new classes of disease processes.

GENOME-WIDE GENE EXPRESSION ANALYSES IN VILI

VILI is a complex pathophysiologic process involving a variety of molecular pathways (*see* recent review in References 3 and 23). This complexity has made a complete elucidation of the molecular mechanisms underlying VILI elusive. There remains considerable controversy regarding the early molecular events induced by mechanical ventilation that initiate alveolar, interstitial, and endothelial injury and the degree to which mechanical ventilation alone is capable of causing inflammation and lung injury (3, 23). To some extent, our limited knowledge of the pathogenesis of ALI and VILI derives from the fact that *a priori* hypotheses based on a single mediator or molecular process are unlikely to adequately describe the multidimensional nature of injury to an organ as complex as the lung. Microarray-based analyses of genome-wide gene expression have the potential to provide an unbiased view of multiple molecular pathways simultaneously, allowing an integrated view of VILI pathogenesis. Recent literature has seen the application of these microarray-based approaches to animal models of VILI which have begun to provide some new insights on VILI pathogenesis. This review will summarize the findings from the four unique microarray expression datasets that have been published to date (24–27) as well as follow-up to this work and related *in vitro* studies.

VILI INDUCED BY MECHANICAL VENTILATION ALONE

The first study to describe microarray-based gene expression profiling of VILI was reported by Copland and colleagues (24). They employed a rat model of VILI to find differentially regulated genes in response to high tidal volume ventilation ($V_T = 5$ ml/kg versus $V_T = 25$ ml/kg, positive end-expiratory pressure = 0 cm/H₂O) for a short period (30 min). That this protocol represents a reasonable model of VILI was shown by the marked decrease in lung compliance and increase in alveolar, interstitial,

and perivascular edema and inflammatory cell infiltrates seen in the animals ventilated with higher tidal volumes.

The microarray analysis in this study was performed using an early generation low-density nylon membrane-based array that contained cDNA probes for less than 2,500 genes. Samples of total RNA purified from the lungs of animals from each experimental group ($n = 4$ per group) were pooled and triplicate arrays were performed for each pool. To determine the genes differentially regulated between the two groups, the authors used the significance analysis of microarrays (SAM) program (28), which uses an error-weighted *t* test to compare expression values. Their criteria for “differentially expressed” included a fold difference ≥ 2 and the derived SAM statistic $\Delta \geq 0.8$ which, according to the authors, roughly translated to an FDR of 1%.

Among the genes that were highly upregulated by high tidal volumes relative to animals ventilated with low tidal volumes were the immediate-early response genes Nur77, Egr1, Btg2, and c-Jun (29, 30). Upregulation of Egr1 suggests activation of protein kinase C (PKC)-mediated pathways and a program of growth factor production (31), while Nur77, Btg2, and c-Jun are implicated in cell death pathways (18, 32, 33), and c-Jun plays an important role in inflammatory processes (34). Other notable findings were upregulation of HSP70, a mediator of cytoprotection (35), and IL-1 β —a gene that, when overexpressed in lung, produces acute lung injury and fibrosis (36), and whose protein product, when elevated in alveolar fluid of patients with ARDS, predicts poor outcome (37). The authors went on to confirm the expression level of these genes using quantitative real-time PCR and localized the expression of a subset of these genes (Egr1, c-Jun, HSP70, and IL-1 β) to the bronchiolar epithelium using *in situ* hybridization. Finally, they demonstrated elevated levels of IL-1 β protein in lungs of mice ventilated with higher tidal volumes.

There are several significant limitations to this study that bear recognition. First, the use of pooling of the biologic replicates calls into question the accuracy of the estimation of the expression level of each gene in each group, particularly if the pool included outliers that skewed the mean. This loss of biological variance information might have led to false conclusions regarding changes in gene expression between groups, although the authors did mitigate this issue, somewhat, by validation work measuring expression levels of a subset of the genes identified using quantitative real-time PCR in individual samples. Second, the use of an array with such a small number of genes ($< 2,500$) makes it difficult to compare the expression profiles obtained in this study with subsequent studies that employed high-density oligonucleotide arrays ($> 12,000$ genes). Third, the use of whole lung tissue as the source of RNA confounds the interpretation of these findings (and findings from several subsequent studies) by the fact that the experimental manipulation (high tidal volume) causes an influx of inflammatory cells. This modifies the cell types that contribute to the mix of RNA purified from whole lung, making it impossible to know which cells were responsible for the change in the measured expression level of a particular gene. This particular issue will be addressed further in descriptions of the study by Dolinay and coworkers below (27). In spite of these limitations, this study provided the first glimpse of the potential utility of array-based expression profiling in VILI and provided some evidence supporting a role for the immediate early gene family in the initial phases of VILI.

VILI WITH MECHANICAL VENTILATION AND INNATE IMMUNE INFLAMMATION

ALI and ARDS can occur in the setting of several risk factors that are characterized by systemic (e.g., sepsis) or local (e.g.,

pneumonia) inflammation (38). Studies in animal models have shown that induction of systemic inflammation with intravenous bacterial lipopolysaccharide (LPS) (26, 39) can cause a synergistic increase in lung injury in the setting of mechanical ventilation. Furthermore, LPS is readily detectable in alveolar lavage fluid from patients with ARDS (40). Two important components of the innate immune pathway recognizing LPS, LPS-binding protein (LBP) and soluble CD14, are elevated in alveolar fluid from patients with ARDS relative to healthy control subjects, and the fluid is capable of facilitating greater inflammatory cytokine production by a macrophage-like cell line (40, 41). These findings suggest that, in ALI and ARDS, VILI could be exacerbated or sustained by the presence of microbial products that induce innate immune inflammatory responses within the alveolar space.

The studies by Altemeier and coworkers and Gharib and colleagues specifically addressed the issue of the interaction between innate immune inflammation and VILI using microarray-based analyses (42, 43). They employed a murine model comparing four groups of animals receiving no mechanical ventilation (MV), noninjurious ($V_T = 10$ ml/kg) MV alone for 4 h, and intratracheal LPS without or with MV (LPS, MV + LPS). In characterizing this model they demonstrated that MV with a tidal volume of 10 ml/kg is minimally injurious over the period studied in terms of alveolar protein leak, alveolar neutrophil concentrations, and inflammatory cytokine production (42). Intratracheal LPS alone caused an increase in neutrophil concentrations, and a modest increase in several inflammatory cytokines, but not an increase in alveolar protein leak. In contrast, intratracheal LPS, in combination with MV, caused a marked increase in protein leak, alveolar neutrophil concentrations, and macrophage inflammatory protein-2 and IL-6 production. Histology corroborated these measurements, demonstrating marked neutrophilic invasion of the alveolar space and alveolar wall thickening. These findings were consistent with a prior study using a similar protocol in rabbits (39).

The authors used a single-color oligonucleotide array platform containing over 22,000 probes measuring the presence of over 12,000 genes. Total RNA was purified from lung homogenates and a separate array was hybridized with each RNA sample ($n = 6$ per group). The intensity values were normalized with the robust multi-chip average algorithm (44) and differential expression of each gene relative to control was determined using the SAM package (28) and multiple hypotheses testing correction to limit the FDR $\leq 1\%$.

Using this approach, the authors found a pattern of gene expression that paralleled the degree of lung injury. The proportion of genes differentially regulated relative to the control group receiving no MV was greatest in the group receiving MV plus intratracheal LPS (1,976 upregulated, 3,753 downregulated) followed by intratracheal LPS alone (1,099 upregulated, 1,832 downregulated) and, finally, MV alone (324 upregulated, 169 downregulated). The authors confirmed a pattern of increasing expression (MV < LPS < MV + LPS) in the genes for CXCL2, CCL3, IL-1 β , IL-6, GADD45- γ , and IRF-7 using quantitative real-time PCR. The finding of GADD45- γ gene up-regulation was confirmed on a protein level using immunohistochemistry showing increased expression of GADD45- γ protein in the alveolar wall of mice in the intratracheal LPS and MV group relative to either condition alone. Grouping of the differentially regulated genes by functional class revealed a preponderance of genes involved in immunity and inflammation (e.g., CCL3, CXCL2, IL-6, IL-1 β), responses to stress (e.g., GADD45- γ), and transcription factors (e.g., IRF7, ATF3). It is notable that GADD45- γ expression had never before been associated with LPS-induced inflammation or VILI before this report, suggesting that microar-

ray-based analyses can identify novel candidate genes in VILI in an unbiased manner. The main limitation to this study is the fact that the authors restricted most of their analyses to genes that were highly upregulated (fold change > 3) in at least one group relative to control. This removed much of the complexity in the expression patterns, including all genes down-regulated by MV or LPS, creating a fairly monotonic view of gene expression in the lung after LPS and/or MV. In addition, this study, similar to that of Copland and coworkers, used whole lung homogenates, making it impossible to tell what proportion of gene expression changes were due to actual changes in transcriptional activity in resident lung cells and how much was due to infiltrating inflammatory cells.

Further analysis of this dataset by Gharib and colleagues (43) using more advanced bioinformatics techniques revealed interesting insights on the potential transcriptional control of the genes differentially expressed under the three experimental conditions. They used software that identified putative transcription factor binding sites that were overrepresented in the promoters of genes found to be differentially regulated by MV and LPS (45). They found that the list of putative transcription factor binding sites overrepresented in the promoter elements of genes upregulated by MV alone (ETF, E2F, NRF1, CREB, and HIF1) was different than those found in the promoter elements of genes upregulated by intratracheal LPS alone (ISRE, cREL, IRF, NFkB, ICSBP, and PU.1). In contrast, genes upregulated by the combination of intratracheal LPS and MV showed a list of putative transcription factor binding sites that represented a fusion of the lists found in either condition alone. This finding complements the findings of their previous publication, suggesting an additive effect of LPS and MV on lung injury and gene expression. While the authors did not confirm that transcription factors actually bound to these putative binding sites, bioinformatics approaches such as this will allow investigators to trace back to the root cause of gene expression changes in this complex animal model.

VILI ANALYZED ACROSS SPECIES AND MODELS

Grigoryev and coworkers have used a different study design to identify genes important to the pathophysiology of VILI (25). The authors have looked for common gene expression signatures in models of VILI performed in four different species (rat, mouse, canine, and human) (25). The rationale behind this approach was that it would identify gene expression responses to stretch that are conserved between species, theoretically enhancing biologically important gene expression signatures. They identified orthologous genes (genes that share a very high degree of sequence similarity between species) across the four species and analyzed the expression of these genes in RNA samples from lung tissue homogenates using oligonucleotide microarrays. The rat and mouse models of VILI employed mechanical ventilation and high tidal volumes for 2 h (mouse) or 4 h (rat). The canine model employed saline lavage-induced VILI of 5 h duration. Finally, they determined gene expression profiles in human pulmonary arterial endothelial cells that were exposed to mechanical stretch for 48 h. Of note, since no commercial array for canine expression analysis existed, the authors relied on the high degree of sequence similarity between humans and canines and used a human oligonucleotide array (Affymetrix U133A; Affymetrix, Santa Clara, CA) to profile the canine RNA samples. After identifying orthologs for 2,887 human reference genes common to all of the oligonucleotide array platforms using the RESOURCER database (46), the authors used a complex algorithm to pool results for control and ventilated specimens across the four species.

The authors identified 69 genes (61 up-regulated and 8 down-regulated) that were differentially regulated by stretch in the pooled dataset. Functional annotation of the 69 genes revealed an overrepresentation of genes involved in immunity, inflammation, regulation of proliferation and cell cycle, coagulation, antimicrobial response, apoptosis, cell-cell signaling, and chemotaxis. Literature searches in the PubMed database showed that 12 of the 69 genes were previously linked with lung injury, including IL-1 β , IL-6, PAI-I, and AQP-1. The authors claimed that this approach looking for evolutionarily conserved responses to VILI was more powerful than studying VILI in an individual species, citing the higher number of genes found to be "significantly" differentially regulated. However, this might have merely reflected the increase in sample size that occurred when the samples obtained from the different species were combined. It remains to be tested whether the cross-species approach is truly a more powerful way to identify important genes in VILI pathogenesis compared with experiments of comparable sample size in a single species. In addition, it should be noted that this approach also used heterogeneous lung tissue as the source of RNA, again muddying the interpretation of the source for apparent changes in gene expression.

IDENTIFYING STRETCH-INDUCED CHANGES IN GENE EXPRESSION SPECIFIC TO LUNG CELLS

As mentioned above, interpretation of changes in gene transcript abundance in RNA samples isolated from whole lung tissue is confounded by changes in cell populations during VILI. Thus, an increase or decrease in the abundance of a transcript could either be due to changes in the lung parenchyma itself or the influx of inflammatory cells such as neutrophils. One specific example of this potential problem is shown by the observation that the gene for S100A9, also known as calgranulin B, was upregulated in all of the *in vivo* VILI models outlined above (24, 25, 42). The product of this gene is a cytosolic protein that constitutes almost half of the protein mass within circulating neutrophils (47), but is also expressed in resident macrophages in the setting of acute inflammation (48). Thus, differential expression of S100A9 could simply be due to an increase in the presence of neutrophils or a change in the expression program of resident alveolar macrophages.

One way to address this potential ambiguity around the source of gene expression changes associated with VILI is to study a more pure population of cells. The study performed by Dolinay and colleagues has helped to clarify the role of cells intrinsic to the lung in the gene expression changes seen in response to high tidal volume mechanical ventilation (27). The authors employed an isolated, perfused mouse lung model using a negative pressure chamber to drive ventilation of the lung. The isolated lungs were perfused with cell-free culture media, *ex vivo*, eliminating peripheral leukocytes from the pulmonary system and allowing the determination of lung-specific changes in gene expression in response to overdistention. Four groups were compared: lungs ventilated with -10 cm H $_2$ O or -25 cm H $_2$ O of end-inspiratory pressure for 3 h, in the presence or absence of LPS in the perfusate. The authors clearly stated that this was a model of overdistention and not VILI *per se*. Consistent with this assertion, the model induced minimal histologic changes and alveolar/interstitial edema was not measured. However, they did show that ventilation with the more negative end-inspiratory pressure (-25 cm H $_2$ O) resulted in higher levels of IL-6 detectable in the perfusate and still higher IL-6 production with addition of LPS to the perfusate. This result parallels the higher concentrations of IL-6 observed in the blood of patients ventilated with injurious tidal volumes (2).

Measurements of genome-wide expression in RNA purified from the isolated, perfused lungs were performed using an oligonucleotide array containing 10,500 mouse genes. There were three to five biologic replicates per group, and differences in gene expression between groups were determined using the SAM package with FDR ≤ 0.1 . Expression profiles revealed 27 genes that were differentially regulated by overstretch alone, including IL-1 β , which was identified as up-regulated in the studies of both Altmeier and coworkers and Copland and colleagues. Grouping these genes by cellular function revealed an overrepresentation of genes of the immune response, receptor binding, mitochondrion, ribosome, growth factor family, apoptosis, signal transduction, and cytokine activity. Of note, the S100A9 (calgranulin B) gene mentioned above was not differentially regulated in this model, suggesting that infiltrating neutrophils indeed drove the up-regulation of this gene measured in other models of VILI using lungs *in situ*. While direct comparisons between this model and those of VILI *in situ* are problematic, given the lack of strong biochemical or histologic evidence of injury, this study does give us the first glimpse of the gene expression changes that derive from resident lung tissue subjected to overdistention.

One study has examined changes in gene expression profiles in a single lung cell type in response to cyclic stretch. Dos Santos et al studied the effect of cyclic stretch (20% elongation, 1 or 4 h duration) on gene expression in the A549 alveolar epithelial cell line in the presence or absence of exogenous TNF- α (49). The authors identified 40 genes that were differentially regulated by TNF- α treatment alone, whereas cyclic stretch alone caused no significant differences in gene expression. The combination of cyclic stretch with TNF- α resulted in augmented expression of 16 genes, including the chemokine CCL20, which is thought to be involved in immune responses within the lung (50). This interaction suggests that the presence of TNF- α , or other inflammatory cytokines, are important for the full pathophysiologic response of the alveolar epithelium to cyclic stretch and provides a potential explanation for the observations, *in vivo*, that co-administration of an inflammatory stimulus such as LPS with mechanical ventilation augments both the degree of lung injury and the scale of the differential gene expression observed (42). This also mirrors the clinical experience that ventilation of patients with no pulmonary or systemic inflammatory process do not manifest signs of lung injury (51, 52).

COMMON THREADS

Observation of a similar pattern of differential gene expression across similar but independent studies of a disease process provides evidence that these shared expression patterns are truly associated with development of the diseases. The four studies described above that used whole lung tissue as a source of RNA were all performed under unique experimental conditions using different species and array platforms. However, the statistical approaches used were relatively similar, with protections against excessive false positive results. Comparison of genes found to be differentially regulated by mechanical ventilation across the four models reveals several genes with similar patterns of regulation, an event that is unlikely to occur by chance. The expression patterns of these genes could provide insight on the mechanisms of VILI. Figure 1 summarizes the 37 genes reported as differentially regulated in two or more of the publications and yields several important observations. First, the highest degree of similarity was seen between the studies by Grigoryev and coworkers (25) and Altmeier and colleagues (42). This could reflect the similarities of the models and the larger sample sizes, leading to more accurate estimations of true gene expression differences.

Gene	Description	Study			
		Grigoryev ¹	Altemeier ²	Copland ³	Dolinay ⁴
AREG	amphiregulin (schwannoma-derived growth factor)				
ARG2	arginase, type II				
ARPC4	actin related protein 2/3 complex, subunit 4, 20kDa				
ATF3	activating transcription factor 3				
BHLHB2	basic helix-loop-helix domain containing, class B, 2				
BTG1	B-cell translocation gene 1, anti-proliferative				
BTG2	BTG family, member 2				
CCL2	chemokine (C-C motif) ligand 2				
CDKN1A	cyclin-dependent kinase inhibitor 1A (p21, Cip1)				
CEBPB	CCAAT/enhancer binding protein (C/EBP), beta				
CEBPD	CCAAT/enhancer binding protein (C/EBP), delta				
CXCR4	chemokine (C-X-C motif) receptor 4				
CYCS	cytochrome c, somatic				
DNAJA1	DnaJ (Hsp40) homolog, subfamily A, member 1				
EGFR	epidermal growth factor receptor				
EIF2S1	eukaryotic translation initiation factor 2, subunit 1 alpha, 35kDa				
F3	coagulation factor III (thromboplastin, tissue factor)				
GADD45A	growth arrest and DNA-damage-inducible, alpha				
GBP2	guanylate binding protein 2, interferon-inducible				
GCLC	glutamate-cysteine ligase, catalytic subunit				
GJA1	gap junction protein, alpha 1, 43kDa (connexin 43)				
HSPA8	heat shock 70kDa protein 8				
IFRD1	interferon-related developmental regulator 1				
IL1B	interleukin 1, beta				
IL1R2	interleukin 1 receptor, type II				
IL6	interleukin 6 (interferon, beta 2)				
LGALS3	lectin, galactoside-binding, soluble, 3 (galectin 3)				
NR4A1	nuclear receptor subfamily 4, group A, member 1 (Nur77)				
PLAUR	plasminogen activator, urokinase receptor				
PTGS2	prostaglandin G/H synthase & cyclooxygenase				
S100A9	S100 calcium binding protein A9 (calgranulin B)				
SERPINE1	serine proteinase inhibitor E1 (plasminogen activator inhibitor -1)				
YWHAZ	Tyr 3 /Trp 5-monooxygenase activation protein, zeta				
AQP1	aquaporin 1 (channel-forming integral protein, 28kDa)				
ADMR	adrenomedullin receptor				
TCF21	transcription factor 21				
TFF2	trefoil factor 2 (spasmolytic protein 1)				

Figure 1. Summary of differentially regulated genes. Genes grouped based on differential regulation in the same direction: *upper panel*, up-regulated; *lower panel*, down-regulated. ¹Genes differentially regulated across murine, rat, canine, and human models of VILI ($p < 0.05$; *black boxes*) (25). ²Genes differentially regulated (95% confidence interval for fold-change not containing 1.0) in murine model of mechanical ventilation alone (*black boxes*) or mechanical ventilation plus intratracheal lipopolysaccharide (*grey boxes*) (26). ³Genes differentially regulated in rat model of mechanical ventilation (fold change ≥ 2 and $D \geq 0.8$) (24). ⁴Genes differentially regulated in murine isolated perfused lung model of mechanical ventilation (false discovery rate ≤ 0.1 ; *black boxes*), or fold change over control ≥ 2.0 (*grey boxes*) (27).

The study by Copland and coworkers (24) had the least number of similarly differentially regulated genes. This is likely to be due to the fact that many of the genes identified in the three other studies were not present on the low-density rat cDNA array ($< 2,500$ genes represented) used for this study.

One gene, IL-1 β , was differentially up-regulated in all four models of VILI. IL-1 β has been shown to be capable of inducing lung injury in rodents (36) and is found in high concentrations in bronchoalveolar lavage fluid from patients with ARDS (37, 53). The fact that IL-1 β was strongly upregulated in the study by Dolinay and colleagues (27) suggests that at least part of the differential gene expression seen in the models testing VILI *in situ* is occurring at the level of the lung parenchyma. Another

consistent finding across the four studies was the upregulation of heat shock protein 70 (HSP70)-related genes. HSPA8 (heat shock 70 kD, protein 8), was found to be upregulated in the Grigoryev, Altemeier, and Dolinay studies. The study by Copland and associates reported upregulation of HSP70, but did not specify which member of the HSP70 family. These findings suggest that HSPA8 might be important to the pathophysiology of VILI.

Characterization of the genes differentially regulated by VILI across the different studies by functional categories defined by the Gene Ontology Database (<http://www.geneontology.org/>) (54) demonstrates some common themes. Table 1 shows the functional categories and the number of genes contained within

TABLE 1. OVERREPRESENTED FUNCTIONAL CATEGORIES

Category	Number of Genes	EASE Score	Fisher Exact
Defense response	12	0.000005	0.000001
Response to biotic stimulus	12	0.000013	0.000003
Immune response	11	0.000015	0.000003
Organismal physiological process	15	0.000020	0.000006
Response to stimulus	15	0.000024	0.000007
Response to external stimulus	13	0.000104	0.000030
Inflammatory response	6	0.000147	0.000013
Innate immune response	6	0.000177	0.000016
Response to stress	10	0.000253	0.000058
Response to pest/pathogen/parasite	8	0.000254	0.000042
Negative regulation of cell proliferation	5	0.000646	0.000052
Response to wounding	6	0.000840	0.000107
Regulation of cell proliferation	5	0.006330	0.000952
Signal transducer activity	12	0.007020	0.003020
Extracellular space	5	0.007660	0.001250

Definition of abbreviation: EASE, Expression Analysis Systematic Explorer.

each category, ranked by the degree of overrepresentation in the category as determined by EASE (Expression Analysis Systematic Explorer) (55) implemented in DAVID (Database for Annotation, Visualization and Integrated Discovery, <http://david.abcc.ncifcrf.gov/tools.jsp>) (55). The most clearly overrepresented theme was responses to pathogens including defense response, response to biotic stimulus, innate immune response, and response to pest/pathogen/parasite (Table 1). Given that the genes selected for this comparison were identified as differentially regulated in the presence of mechanical ventilation alone, this finding suggests that the response to alveolar overdistention might mimic an innate immune inflammatory response. This would help to explain why co-administration of an innate immune stimulus, such as LPS, with mechanical ventilation can result in an additive increase in lung injury. The other functional categories overrepresented are related to cell proliferation and repair, including negative regulation of cell proliferation, response to wounding, regulation of cell proliferation, and extracellular space (Table 1). The genes in these categories may reflect the initiation of the fibroproliferative process that, in human ARDS, has been associated with poor outcomes (56–58). Overall, the functional categories identified are biologically plausible, which suggests that microarray-based analysis of VILI provide valid results. Of note, a search of the PubMed database using PubMatrix (<http://pubmatrix.grc.nia.nih.gov/>) (59) reveals that less than half of the genes in Figure 1 have been previously associated with lung injury in the literature, again suggesting that microarray-based analysis is capable of revealing novel candidate genes in the pathogenesis of VILI. Further studies are needed to validate this assertion.

FUTURE USES OF GENOME-WIDE GENE EXPRESSION ANALYSES IN VILI

The studies described above highlight the potential utility of microarray-based analyses in the study of VILI and have provided novel insight on its pathogenesis. The genes identified could have direct functional significance for VILI reflecting functional components upstream of the gene, or simply an epiphenomenon. Future studies will need to differentiate these possibilities using further work in animal models and cell culture systems. Investigators can also establish relevance in human ALI using genetic approaches. For instance, Ye and coworkers found that the gene encoding pre-B cell enhancing factor (PBEF) was up-regulated in a canine model of VILI (60). Using a human poly-

morphic allele that increases the expression level of the PBEF gene, they demonstrated that patients carrying the high expression allele were at much higher risk of developing ALI (60), highly suggestive of a causative role for PBEF in ALI pathogenesis. Taken a step further, one could imagine clinical studies in which genome-wide genotyping and gene expression are performed in parallel with collection of clinical phenotypes. A genotype associated with the clinical phenotype could be quickly assigned functional status through the identification of a proximal gene whose expression level is correlated with the genotype. This sort of approach has been used in animal models of diabetes with some success (61). Recent advances in our understanding of genetic variation within humans and development of inexpensive high-throughput array-based genotyping platforms should facilitate such studies.

A key question will be what cell type to use for gene expression profiling in human studies of VILI. Given that lung biopsies in the setting of critical illness are unlikely to be feasible for any reasonably powered study, an alternative might be alveolar macrophages. These cells are easily obtained via bronchoalveolar lavage and are likely to provide a transcriptional “window” into the state of the alveolar space. Bronchoalveolar lavage-derived cells have been used for gene expression profiling in smokers (62), scleroderma (63), and acute lung transplant rejection (63), and have yielded promising results.

Finally, more advanced study designs will provide greater resolution to discern mechanisms in human and animal studies of VILI. In particular, incorporation of a greater number of time points will allow a longitudinal view of gene expression patterns. Longitudinal analysis of genome-wide gene expression has been used successfully in the study of atherosclerosis (64) and will provide richer datasets from which to develop gene interaction networks that might underlie the pathogenesis of VILI.

Conflict of Interest Statement: M.M.W. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

References

- Webb HH, Tierney DF. Experimental pulmonary edema due to intermittent positive pressure ventilation with high inflation pressures: protection by positive end-expiratory pressure. *Am Rev Respir Dis* 1974;110:556–565.
- The Acute Respiratory Distress Syndrome Network. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *N Engl J Med* 2000;342:1301–1308.
- Tremblay LN, Slutsky AS. Ventilator-induced lung injury: from the bench to the bedside. *Intensive Care Med* 2006;32:24–33.
- DeRisi J, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, Chen Y, Su YA, Trent JM. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat Genet* 1996;14:457–460.
- Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995;270:467–470.
- Lipshutz RJ, Morris D, Chee M, Hubbell E, Kozal MJ, Shah N, Shen N, Yang R, Fodor SP. Using oligonucleotide probe arrays to access genetic diversity. *Biotechniques* 1995;19:442–447.
- Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet* 2006;7:55–65.
- Quackenbush J. Microarray analysis and tumor classification. *N Engl J Med* 2006;354:2463–2472.
- van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;347:1999–2009.
- Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;286:531–537.
- Verhaak RG, Goudswaard CS, van Putten W, Bijl MA, Sanders MA, Hagens W, Uitterlinden AG, Erpelinck CA, Delwel R, Lowenberg B,

- et al.* Mutations in nucleophosmin (NPM1) in acute myeloid leukemia (AML): association with other gene abnormalities and previously established gene expression signatures and their favorable prognostic significance. *Blood* 2005;106:3747–3754.
12. Beer DG, Kardia SL, Huang CC, Giordano TJ, Levin AM, Misek DE, Lin L, Chen G, Gharib TG, Thomas DG, *et al.* Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 2002; 8:816–824.
 13. Kuperman DA, Lewis CC, Woodruff PG, Rodriguez MW, Yang YH, Dolganov GM, Fahy JV, Erle DJ. Dissecting asthma using focused transgenic modeling and functional genomics. *J Allergy Clin Immunol* 2005;116:305–311.
 14. Pardo A, Gibson K, Cisneros J, Richards TJ, Yang Y, Becerril C, Yousem S, Herrera I, Ruiz V, Selman M, *et al.* Up-regulation and profibrotic role of osteopontin in human idiopathic pulmonary fibrosis. *PLoS Med* 2005;2:e251.
 15. Bull TM, Coldren CD, Moore M, Sotto-Santiago SM, Pham DV, Nana-Sinkam SP, Voelkel NF, Geraci MW. Gene microarray analysis of peripheral blood cells in pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2004;170:911–919.
 16. de Reynies A, Geromin D, Cayuela JM, Petel F, Dessen P, Sigaux F, Rickman DS. Comparison of the latest commercial short and long oligonucleotide microarray technologies. *BMC Genomics* 2006;7:51.
 17. Larkin JE, Frank BC, Gavras H, Sultana R, Quackenbush J. Independence and reproducibility across microarray platforms. *Nat Methods* 2005;2:337–344.
 18. Winoto A, Littman DR. Nuclear hormone receptors in T lymphocytes. *Cell* 2002;109:S57–S66.
 19. Tian Q, Stepaniants SB, Mao M, Weng L, Feetham MC, Doyle MJ, Yi EC, Dai H, Thorsson V, Eng J, *et al.* Integrated genomic and proteomic analyses of gene expression in Mammalian cells. *Mol Cell Proteomics* 2004;3:960–969.
 20. Bakay M, Chen YW, Borup R, Zhao P, Nagaraju K, Hoffman EP. Sources of variability and effect of experimental approach on expression profiling data interpretation. *BMC Bioinformatics* 2002;3:4.
 21. Novak JP, Sladek R, Hudson TJ. Characterization of variability in large-scale gene expression data: implications for study design. *Genomics* 2002;79:104–113.
 22. Curran-Everett D. Multiple comparisons: philosophies and illustrations. *Am J Physiol Regul Integr Comp Physiol* 2000;279:R1–R8.
 23. Frank JA, Matthay MA. Science review: mechanisms of ventilator-induced injury. *Crit Care* 2003;7:233–241.
 24. Copland IB, Kavanagh BP, Engelberts D, McKerlie C, Belik J, Post M. Early changes in lung gene expression due to high tidal volume. *Am J Respir Crit Care Med* 2003;168:1051–1059.
 25. Grigoryev DN, Ma SF, Irizarry RA, Ye SQ, Quackenbush J, Garcia JG. Orthologous gene-expression profiling in multi-species models: search for candidate genes. *Genome Biol* 2004;5:R34.
 26. Altmeier WA, Matute-Bello G, Frevert CW, Kawata Y, Kajikawa O, Martin TR, Glenny RW. Mechanical ventilation with moderate tidal volumes synergistically increases lung cytokine response to systemic endotoxin. *Am J Physiol Lung Cell Mol Physiol* 2004;287:L533–L542.
 27. Dolinay T, Kaminski N, Felgendreher M, Kim HP, Reynolds P, Watkins SC, Karp D, Uhlig S, Choi AM. Gene expression profiling of target genes in ventilator-induced lung injury. *Physiol Genomics* 2006;26: 68–75.
 28. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 2001;98:5116–5121.
 29. Hazel TG, Nathans D, Lau LF. A gene inducible by serum growth factors encodes a member of the steroid and thyroid hormone receptor superfamily. *Proc Natl Acad Sci USA* 1988;85:8444–8448.
 30. Herschman HR. Primary response genes induced by growth factors and tumor promoters. *Annu Rev Biochem* 1991;60:281–319.
 31. Silverman ES, Collins T. Pathways of Egr-1-mediated gene transcription in vascular biology. *Am J Pathol* 1999;154:665–670.
 32. Tirone F. The gene PC3/TIS21/BTG2), prototype member of the PC3/BTG/TOB family: regulator in control of cell growth, differentiation, and DNA repair? *J Cell Physiol* 2001;187:155–165.
 33. Dunn C, Wiltshire C, MacLaren A, Gillespie DA. Molecular mechanism and biological functions of c-Jun N-terminal kinase signalling via the c-Jun transcription factor. *Cell Signal* 2002;14:585–593.
 34. Arndt PG, Young SK, Lieber JG, Fessler MB, Nick JA, Worthen GS. Inhibition of c-Jun N-terminal kinase limits lipopolysaccharide-induced pulmonary neutrophil influx. *Am J Respir Crit Care Med* 2005; 171:978–986.
 35. Beere HM, Green DR. Stress management: heat shock protein-70 and the regulation of apoptosis. *Trends Cell Biol* 2001;11:6–10.
 36. Kolb M, Margetts PJ, Anthony DC, Pitossi F, Gauldie J. Transient expression of IL-1 β induces acute lung injury and chronic repair leading to pulmonary fibrosis. *J Clin Invest* 2001;107:1529–1536.
 37. Meduri GU, Kohler G, Headley S, Tolley E, Stentz F, Postlethwaite A. Inflammatory cytokines in the BAL of patients with ARDS: persistent elevation over time predicts poor outcome. *Chest* 1995;108:1303–1314.
 38. Hudson LD, Milberg JA, Anardi D, Maunder RJ. Clinical risks for development of the acute respiratory distress syndrome. *Am J Respir Crit Care Med* 1995;151:293–301.
 39. Bregeon F, Delpierre S, Chetaille B, Kajikawa O, Martin TR, Auttilo-Touati A, Jammes Y, Pugin J. Mechanical ventilation affects lung function and cytokine production in an experimental model of endotoxemia. *Anesthesiology* 2005;102:331–339.
 40. Martin TR, Rubenfeld GD, Ruzinski JT, Goodman RB, Steinberg KP, Leturcq DJ, Moriarty AM, Raghu G, Baughman RP, Hudson LD. Relationship between soluble CD14, lipopolysaccharide binding protein, and the alveolar inflammatory response in patients with acute respiratory distress syndrome. *Am J Respir Crit Care Med* 1997;155: 937–944.
 41. Wurfel MM, Radella F, Ruzinski JT, Park DR, Steinberg KP, Martin TR. Bronchoalveolar lavage fluid from patients with acute respiratory distress syndrome facilitates bacterial lipopolysaccharide recognition: Role for lipopolysaccharide-binding protein. *Am J Respir Crit Care Med* 2003;167:A760.
 42. Altmeier WA, Matute-Bello G, Gharib SA, Glenny RW, Martin TR, Liles WC. Modulation of lipopolysaccharide-induced gene transcription and promotion of lung injury by mechanical ventilation. *J Immunol* 2005;175:3369–3376.
 43. Gharib SA, Liles WC, Matute-Bello G, Glenny RW, Martin TR, Altmeier WA. Computational identification of key biological modules and transcription factors in acute lung injury. *Am J Respir Crit Care Med* 2006;173:653–658.
 44. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 2003;19:185–193.
 45. Elkon R, Linhart C, Sharan R, Shamir R, Shiloh Y. Genome-wide in silico identification of transcriptional regulators controlling the cell cycle in human cells. *Genome Res* 2003;13:773–780.
 46. Tsai J, Sultana R, Lee Y, Pertege G, Karamycheva S, Antonescu V, Cho J, Parvizi B, Cheung F, Quackenbush J. RESOURCERER: a database for annotating and linking microarray resources within and across species. *Genome Biol* 2001;2:SOFTWARE0002.
 47. Hessian PA, Edgeworth J, Hogg N. MRP-8 and MRP-14, two abundant Ca(2+)-binding proteins of neutrophils and monocytes. *J Leukoc Biol* 1993;53:197–204.
 48. Hu SP, Harrison C, Xu K, Cornish CJ, Geczy CL. Induction of the chemotactic S100 protein, CP-10, in monocyte/macrophages by lipopolysaccharide. *Blood* 1996;87:3919–3928.
 49. dos Santos CC, Han B, Andrade CF, Bai X, Uhlig S, Hubmayr R, Tsang M, Lodyga M, Keshavjee S, Slutsky AS, *et al.* DNA microarray analysis of gene expression in alveolar epithelial cells in response to TNF α , LPS, and cyclic stretch. *Physiol Genomics* 2004;19:331–342.
 50. Lin TJ, Maher LH, Gomi K, McCurdy JD, Garduno R, Marshall JS. Selective early production of CCL20, or macrophage inflammatory protein 3 α , by human mast cells in response to *Pseudomonas aeruginosa*. *Infect Immun* 2003;71:365–373.
 51. Pelosi P, Ravagnan I, Giurati G, Panigada M, Bottino N, Tredici S, Echer G, Gattinoni L. Positive end-expiratory pressure improves respiratory function in obese but not in normal subjects during anesthesia and paralysis. *Anesthesiology* 1999;91:1221–1231.
 52. Wrigge H, Uhlig U, Zinserling J, Behrends-Callsen E, Ottersbach G, Fischer M, Uhlig S, Putensen C. The effects of different ventilatory settings on pulmonary and systemic inflammatory responses during major surgery. *Anesth Analg* 2004;98:775–781. (table of contents.).
 53. Pugin J, Ricou B, Steinberg KP, Suter PM, Martin TR. Proinflammatory activity in bronchoalveolar lavage fluids from patients with ARDS, a prominent role for interleukin-1. *Am J Respir Crit Care Med* 1996;153: 1850–1856.
 54. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000;25:25–29.
 55. Hosack DA, Dennis G Jr, Sherman BT, Lane HC, Lempicki RA. Identifying biological themes within lists of genes with EASE. *Genome Biol* 2003;4:R70.

56. Clark JG, Milberg JA, Steinberg KP, Hudson LD. Elevated lavage levels of N-terminal peptide of type III procollagen are associated with increased fatality in adult respiratory distress syndrome. *Chest* 1994; 105:126S–127S.
57. Clark JG, Milberg JA, Steinberg KP, Hudson LD. Type III procollagen peptide in the adult respiratory distress syndrome. Association of increased peptide levels in bronchoalveolar lavage fluid with increased risk for death. *Ann Intern Med* 1995;122:17–23.
58. Steinberg KP, Hudson LD, Goodman RB, Hough CL, Lanken PN, Hyzy R, Thompson BT, Ancukiewicz M. Efficacy and safety of corticosteroids for persistent acute respiratory distress syndrome. *N Engl J Med* 2006;354:1671–1684.
59. Becker KG, Hosack DA, Dennis G Jr, Lempicki RA, Bright TJ, Cheadle C, Engel J. PubMatrix: a tool for multiplex literature mining. *BMC Bioinformatics* 2003;4:61.
60. Ye SQ, Simon BA, Maloney JP, Zambelli-Weiner A, Gao L, Grant A, Easley RB, McVerry BJ, Tudor RM, Standiford T, *et al.* Pre-B-cell colony-enhancing factor as a potential novel biomarker in acute lung injury. *Am J Respir Crit Care Med* 2005;171:361–370.
61. Mehrabian M, Allayee H, Stockton J, Lum PY, Drake TA, Castellani LW, Suh M, Armour C, Edwards S, Lamb J, *et al.* Integrating genotypic and expression data in a segregating mouse population to identify 5-lipoxygenase as a susceptibility gene for obesity and bone traits. *Nat Genet* 2005;37:1224–1233.
62. Woodruff PG, Koth LL, Yang YH, Rodriguez MW, Favoreto S, Dolganov GM, Paquet AC, Erle DJ. A distinctive alveolar macrophage activation state induced by cigarette smoking. *Am J Respir Crit Care Med* 2005;172:1383–1392.
63. Luzina IG, Atamas SP, Wise R, Wigley FM, Xiao HQ, White B. Gene expression in bronchoalveolar lavage cells from scleroderma patients. *Am J Respir Cell Mol Biol* 2002;26:549–557.
64. Tabibiazar R, Wagner RA, Ashley EA, King JY, Ferrara R, Spin JM, Sanan DA, Narasimhan B, Tibshirani R, Tsao PS, *et al.* Signature patterns of gene expression in mouse atherosclerosis and their correlation to human coronary disease. *Physiol Genomics* 2005;22:213–226.