Increased Expression Of Neutrophil-Related Genes In Patients With Early Sepsis-Induced ARDS

Kirsten Neudoerffer Kangelaris MD,1 Arun Prakash, MD PhD,2 Kathleen D. Liu MD PhD,2,3 Bradley Aouizerat PhD, MAS4,5 Prescott G. Woodruff MD, MPH2,6 David J. Erle MD,6 Angela Rogers MD,6 Eric J. Seeley MD,6 Jeffrey Chu BS,2 Tom Liu BS,2 Thomas Osterberg-Deiss BS,2 Hanjing Zhuo MPH,2 Michael A. Matthay MD,2,3,6 Carolyn S. Calfee MD MAS2,3,6

University of California–San Francisco, San Francisco, California
1Department of Medicine, Division of Hospital Medicine
2Departments of Medicine and Anesthesia
3Cardiovascular Research Institute,
4Department of Physiologic Nursing
5Institute for Human Genetics
6Department of Medicine, Division of Pulmonary, Critical Care, Allergy and Sleep Medicine
7Department of Pulmonary and Critical Care
8Stanford University
9Department of Pulmonary and Critical Care

Corresponding author:
Kirsten N. Kangelaris
Box 0131, 533 Parnassus Ave UC Hall, University of California, San Francisco, San Francisco, CA 94143-0131.
E-mail: kkangelaris@medicine.ucsf.edu
Telephone: (415) 476-4852; Fax: (415) 476-4818

Funding sources: At the time the research was conducted Dr. Kangelaris was supported by the NIH National Center for Advancing Translational Sciences through UCSF-CTSI KL2 TR000143, and NHLBI 1K23HL116800-01. Dr. Prakash was supported by a FAER MRTG-BS. Dr. Calfee was supported by NHLBI HL110969. Dr. Matthay was supported by NHLBI R37HL51856.

Author’s contributions:
KNK, CSC and MAM contributed to the conception, hypothesis delineation, design of study, acquisition or analysis, and interpretation of data, and drafting and revising the article
PGW, EJS and AR, DJE contributed to the conception, hypothesis delineation, interpretation of data and were involved in revision of the article prior to submission
AP, BA, KDL, JC, HZ, TL, TO-D were involved in the acquisition of the data, analysis of data and in the revision of the article prior to submission.
All authors were involved in revising the manuscript critically for important intellectual content.
All authors approved the version to be published.

Conflicts of interest: No author reports a conflict of interest. Dr. Calfee has served on medical advisory boards for Glaxo Smith Kline and received grant funding from Glaxo Smith Kline.
ABSTRACT

Background: The early sequence of events leading to the development of the acute respiratory distress syndrome (ARDS) in patients with sepsis remains inadequately understood. The purpose of this study was to identify changes in gene expression early in the course of illness, when mechanisms of injury may provide the most relevant treatment and prognostic targets.

Methods: We collected whole blood RNA in critically ill patients admitted from the Emergency Department to the intensive care unit within 24 hours of admission at a tertiary care center. Whole genome expression was compared in patients with sepsis and ARDS to patients with sepsis alone. We selected genes with > 1 log₂ fold change and false discovery rate <0.25, determined their significance in the literature and performed pathway analysis.

Results: Several genes were upregulated in 29 patients with sepsis with ARDS compared to 28 patients with sepsis alone. The most differentially expressed genes included key mediators of the initial neutrophil response to infection: olfactomedin 4 (OLFM4), lipocalin 2 (LCN2), CD24, and bactericidal/permeability-increasing protein (BPI). These gene expression differences withstood adjustment for age, sex, study batch, white blood cell count and presence of pneumonia or aspiration. Pathway analysis demonstrated over-representation of genes involved in known respiratory and infection pathways.

Conclusions: These data indicate that several neutrophil-related pathways may be involved in the early pathogenesis of sepsis-related ARDS. In addition, identifiable gene expression
differences occurring early in the course of sepsis-related ARDS may further elucidate understanding of the neutrophil-related mechanisms in progression to ARDS.
INTRODUCTION

The Acute Respiratory Distress Syndrome (ARDS) is a common and often early complication of sepsis, but why only a fraction of patients with sepsis develop ARDS remains incompletely understood. (54) The National Heart, Lung and Blood Institute has recently created a clinical network designed to enroll patients with sepsis and other predisposing conditions into clinical trials focused on prevention of ARDS and treatment of patients with early acute lung injury before the development of ARDS. (5, 29, 42, 43) Thus, insights into the early mechanisms that lead from the endothelial damage of sepsis to the lung specific capillary leak of ARDS are needed.

One strategy for identifying sepsis patients at higher risk for developing ARDS has been to measure protein biomarkers in the plasma. We and other investigators have reported that this approach may have predictive value, especially for biological markers of endothelial injury. (8, 55) However, this approach is limited to well-described candidate biomarkers. In contrast, whole blood gene expression has the potential to not only validate known biomarkers but for the novel discovery of mediators, pathways and/or biomarkers in the pathogenesis of ARDS. Gene expression has been successful in identifying subgroups of severity in sepsis, (49, 69-71) and is supported by studies demonstrating that gene expression in sepsis varies significantly over time. (57) However, there have been only three clinical microarray studies published in sepsis-related ARDS (19, 23, 35), and these studies focused on patients enrolled later in their course of illness, with as long as a 48 hour delay between admission and the acquisition of blood samples.

Based on the foundational concept that changes in gene transcription precede and guide the majority of important biologic processes, we sought to understand early changes, in our case just hours after presentation to medical care, in whole blood gene transcription that distinguish septic
patients that develop lung injury from those who do not. The primary hypothesis of the current
study was that gene expression in whole blood samples from critically ill human subjects with
early severe sepsis would reveal differential expression of biologically plausible genes in patients
with ARDS compared to patients without ARDS.

MATERIALS AND METHODS

Human subjects. We studied prospectively enrolled critically ill patients admitted to a
tertiary care hospital ICU from the emergency department between October 2009 and April 2012
as part of the ongoing Early Assessment of Renal and Lung Injury (EARLI) cohort, the details of
which have been previously published.(8, 75) In this study, critically ill patients admitted via the
emergency department were enrolled at the time of triage to the ICU. For the current analysis,
we included patients from this cohort who met consensus criteria for severe sepsis and had whole
blood RNA samples drawn within 24 hours of admission to the ICU.(1) Sepsis was defined as
documented or suspected infection in the presence of two or more of the following
characteristics of the systemic inflammatory response syndrome (SIRS): (1) temperature >38°C
or < 36°C; (2) heart rate > 90 beats per minute; (3) respiratory rate > 20 breaths per minute or
PaCO2 < 32 mmHg or need for mechanical ventilation; or (4) white blood cell count >12,000/cu
mm, <4,000/cu mm.(12) Patients with at least one organ dysfunction were classified as severe
sepsis. Hypotension, need for mechanical ventilation, oliguria, altered mental status and lactic
acidosis were examples of organ dysfunction included. Shock was defined as a systolic blood
pressure <90mmHg or vasopressor use. Patients were defined as having ARDS if they met
criteria as defined by the Berlin definition of ARDS within 24 hours of enrollment to the
study.(51) In order to study a clearly defined clinical phenotype, patients were excluded if they
had an equivocal diagnosis of ARDS based on chest radiograph or absent arterial blood gas or if they did not receive positive pressure ventilation (n=10). Patients were also excluded if whole blood suitable for RNA isolation was not collected (Flow diagram, Figure 1). Sample size was determined based on a prior gene expression study of sepsis-related ARDS.(35) Fifty-seven total patients were enrolled; 29 of these had sepsis without lung injury and 28 had sepsis with ARDS. The study was approved by the University of California, San Francisco Institutional Review Board. Procedures for informed consent have been previously described.(8)

**Specimen collection and clinical phenotyping.** Whole blood was collected within 24 hours of ICU admission for the isolation of RNA (see methods below); plasma was also collected simultaneously. Adjudication of sepsis was performed by experienced intensive care physicians. ARDS diagnosis was determined by two board-certified intensive care physicians blinded to the microarray outcome data. Patients were divided into those with and those without ARDS and were followed to hospital discharge or until hospital day 60, whichever came first. Sixty-day mortality was defined as in-hospital death up to day 60 of follow-up.

**Biologic Sample Collection, Processing and Measurements.** Whole blood was stored in coded and de-identified PAXgene™ tubes (PreAnalytiX, Hombrechtikon, CH) tubes at -80°C. RNA was extracted using the PAXgene™ miRNA kit (Qiagen, Valencia, CA) and processed in two separate batches balanced for cases and controls using identical materials and methods. Quality, quantity and purity of the RNA was assessed by electrophoresis (Bioanalyzer, Agilent, Inc) and spectrophotometry. The NuGEN Ovation® Whole Blood Reagent (NuGEN technology, San Carlos, CA) was used for amplification, fragmentation and biotin-labeling, which enables gene expression of whole blood RNA without the need for additional globin reduction.(44) Labeled cDNA was hybridized to the Affymetrix Human GeneChip® Gene 1.0
ST array (Affymetrix, Santa Clara, CA), integrating 28,869 genes with 764,885 probes (~26 probes per gene). Signal intensity fluorescent images were read using the Affymetrix Model 3000 Scanner and converted into GeneChip probe results files (CEL) using Command and Expression Console software (Affymetrix). The experimental design, RNA extraction and microarray processing and bioinformatics were MIAME (minimum information about a microarray experiment)-compliant, and complete raw and normalized microarray data are available through the Gene Expression Omnibus (GEO) of the National Center for Biotechnology Information (GEO accession number GSE66890).(3)

In order to determine if elevated gene expression was corroborated by an increase in the protein product, we measured plasma levels of lipocalin 2. We selected the lipocalin 2 protein, because it is easily measurable in serum samples and has been associated with the pathogenesis of ARDS. (18, 28, 34) Plasma lipocalin 2 was quantified in duplicate and in a blinded fashion using a commercially available two-antibody sandwich ELISA (NGAL a.k.a LCN2 ELISA Kit [Cat.No. KIT 036], Bioporto, Gentofte, Denmark) according to the manufacturer’s instructions.

Quantitative reverse transcription real-time polymerase chain reaction. Differentially expressed genes were selected for qPCR confirmation. TaqMan-specific inventoried primers for olfactomedin 4 (OLFM4), lipocalin 2 (LCN2), bactericidal/permeability-increasing protein (BPI), CD24 molecule (CD24), hydroxycarboxylic acid receptor 3 (HCAR3), and membrane metallo-endopeptidase (MME) were used to measure the message levels of these genes (Life Technologies, Carlsbad, CA). Housekeeping genes glucoronidase beta (GUSB) and hypoxanthine phosphoribosyltransferase 1 (HPRT1) were also measured. Quantitative real-time polymerase chain reaction was performed using the ABI Prism 7000 Sequence Detection System (Life Technologies, Carlsbad, CA). Run method: Polymerase chain reaction activation at 95°C
for 20s was followed by 40 cycles of 1s at 95°C and 20s at 60°C.

The average threshold count (Ct) value of 2–3 technical replicates was used in all calculations. The average Ct values of the internal controls (housekeeping genes) was used to calculate ΔCt values for the samples. Data analysis was performed using the $2^{-\Delta\Delta Ct}$ method, and the data were corrected for statistical analysis using log transformation, mean centering, and autoscaling(46, 56, 67). The lowest value was set to an RQ of 1. Where a message was undetected after 40 cycles of PCR, a Ct value of 40 was assigned since 40 cycles is the limit of detection.

**Statistical Methods.** We compared whole genome gene expression in patients with ARDS and sepsis to those with sepsis alone. Microarrays were normalized for array-specific effects using Affymetrix's "Robust Multi-Array" (RMA) normalization. Normalized array values were reported on a log2 scale (Average normalized expression is typically ~7.0.). For statistical analyses, we removed all array probesets where no experimental groups had an average log2 intensity greater than 3.0. This is a standard cutoff, below which expression is indistinguishable from background noise. Moderated t-statistics, fold-change and statistical significance were calculated for each gene. Linear models were fitted for each gene using the Bioconductor "limma" package in the R statistical environment (www.R-project.org).

We considered batch and relevant clinical differences between the comparison groups in our adjusted analyses. Differences in patient age, sex, RNA processing, and peripheral white blood cell count (WBC) can introduce variability in expression levels.(25, 31, 59, 60) Specifically, the leukemoid reaction characteristic of sepsis may affect the measured expression of leukocyte-derived genes. For example a patient with a white blood cell count of 18,000 cells/ul may have increased expression of some genes simply due to an increase in the number of
neutrophils rather than a functional upregulation of the neutrophil itself. Since pulmonary and non-pulmonary sepsis may lead to ARDS through independent mechanisms, we accounted for indirect (non-pulmonary) versus direct (pneumonia or aspiration) risk factors for lung injury in the analysis.\(^{(14, 61)}\)

Our adjusted analysis included two models: In \textit{Model 1} we adjusted for age, batch number, sex and white blood cell count (WBC); in \textit{Model 2} we adjusted for all characteristics in \textit{Model 1} and for presence of direct lung injury as a risk factor for ARDS. The rationale for adjusting for the presence of direct lung injury in the current study was to assess whether the changes in gene expression were due to more prevalent direct lung injury (pneumonia or aspiration) rather than biologic processes related to ARDS specifically. In a sensitivity analysis, we limited our sample to patients with shock in order to determine whether the differentially expressed genes in ARDS were related to more prevalent shock in the ARDS subgroup.

We selected differentially expressed genes with an \textit{a priori} cutpoint of > 1 log\(_2\) fold change and a false discovery rate (FDR) of < 0.25. The FDR, calculated using the Benjamini-Hochberg method,\(^{(10, 11, 52)}\) accounts for the fact that thousands of genes were tested. The FDR values indicate the expected fraction of falsely declared differentially-expressed genes among the total set of differentially-expressed genes (i.e., FDR = 0.25 would indicate that the result is likely to be valid 3 out of 4 times). In order to determine the biologic plausibility of differentially expressed genes identified, we systematically searched PubMed for relevant basic and clinical studies reporting that the genes or their products were relevant to the pathogenesis of ARDS. Both human and animal research was included. A sample of our search strategy was as follows: "lung injury" OR "respiratory distress syndrome" OR "acute respiratory" OR
pneumonia OR ARDS) AND ("lipocalin 2" OR “neutrophil gelatinase-associated lipocalin” OR “LCN2” OR “LCN 2” OR “NGAL”).

In order to identify signaling pathways and gene networks we included all differentially expressed genes with FDR < 0.25 and log₂ fold change of > 0.2625 using the Ingenuity Pathways Analysis (IPA) application (Ingenuity Systems, Redwood City, CA, USA). The reported p-values are derived using corrections for multiple comparisons and provide an estimate of the probability that a given enrichment is present by chance alone.

The appropriate statistical tests (Spearman, chi-squared, t-test, Mann-Whitney) were used for bivariate analyses of clinical characteristics, qPCR confirmation values and plasma lipocalin 2 and were performed using STATA version 12 (STATA Corp, College Station, TX).

RESULTS

Baseline and Clinical Characteristics

The baseline and clinical characteristics of the study sample according to ARDS diagnosis are described in Table 1. White blood cell count and absolute neutrophil count (ANC) were lower in septic patients with ARDS versus septic patients without ARDS. Direct lung injury (pneumonia or aspiration) as an ARDS risk factor was more prevalent in the ARDS patients (P = 0.05). Septic patients with ARDS were more severely ill than sepsis patients without ARDS with significantly increased APACHE III scores (P = 0.002) and shock (P = 0.02).

Candidate genes identified – Unadjusted and Model 1: adjusted for age, sex, batch and WBC
Of 28,127 genes annotated on the array, 24,403 (97%) met an average log₂ intensity greater than 3 and were analyzed for differential expression. In unadjusted analysis, differentially expressed genes meeting a cut point of > 1 log₂ fold difference and FDR < 0.25 are presented in Table 2. The upregulated genes in septic patients with ARDS included olfactomedin 4 (OLFM4), lipocalin 2 (LCN2), the CD24 molecule (CD24), and bactericidal/permeability-increasing protein (BPI). Hydroxycarboxylic acid receptor 3 (HCAR3) and membrane metalloendopeptidase (MME) were downregulated in septic patients with ARDS. After adjustment for age, sex, batch and the number of white blood cells, the same genes were differentially expressed (Table 3).

Literature search of the PubMed database identified that three of the four identified upregulated genes, LCN2, CD24 and BPI, have been previously described in studies of the pathogenesis of ARDS (Table 3). The downregulated genes identified in our study were not found to have relevant citations to ARDS pathogenesis in the literature.

In a sensitivity analysis limited to the 33 patients with septic shock, we found that the most differentially expressed genes and degree of relative differential expression in unadjusted analysis were similar to those in the whole cohort. The false discovery rates were somewhat increased in this subgroup, likely attributable to decreased sample size (data not shown). Candidate genes identified – Model 2, adjusted for age, sex, batch and WBC, and direct lung injury

Further adjustment of the model to account for the presence of direct lung injury yielded the same candidate genes with similar log₂-fold changes and FDRs identified in Model 1. In addition, there were 7 additional genes (Table 4, shaded in grey) that met the cutpoint of > 1 log₂ fold change and FDR <0.25, not identified in unadjusted analysis or in Model 1. Two of the
seven genes had relevant citations in the literature: matrix metallopeptidase 8 (MMP8), also known as neutrophil collagenase, and haptoglobin (HP), both of which were upregulated in ARDS. Post-hoc analysis of these genes in both unadjusted analysis and Model 1 demonstrated similar log$_2$-fold changes and FDRs that did not quite meet the pre-defined significance cutpoints (data not shown).

**Correlation of candidate genes identified but no association with ANC**

Since several of the candidate genes identified are involved in neutrophil mediated host defense mechanisms, and because elevation in ANC is an important mediator of innate immune response in sepsis,(36) we wanted to determine if the neutrophil-related genes identified (OLFM4, LCN2, CD24, BPI and MMP8) were dependent on number of circulating neutrophils. As shown in Figure 2a and 2b, there was no correlation between the expression levels of the genes and ANC, except for HCAR3, which was positively but weakly correlated with the absolute neutrophil count (Spearman’s rho = 0.41, $P = 0.002$, Figure 2a). In contrast, upregulated genes in Models 1 and 2 were positively correlated with other upregulated genes and demonstrated a negative correlation with downregulated genes (all P-values < 0.05). With the exception of SNORD64, downregulated genes in models 1 and 2 were also significantly intercorrelated (Figure 2a and 2b).

**Validation of differentially expressed genes**

Quantitative PCR, performed for all genes identified in Model 1, was utilized to confirm the differential gene expression identified in the microarray analysis (OLFM4, LCN2, CD24, BPI). The two downregulated genes could not be confirmed by qPCR (MME, HCAR3), as
shown in Figure 3. Normalized array expression and qPCR expression levels of genes were highly correlated (data not shown) (Spearman’s rho > 0.7, P < 0.0001 for all genes).

Plasma levels of lipocalin 2 protein correlated well with the microarray gene expression levels (Spearman’s rho = 0.67, P < 0.0001, Figure 4a). There was also a modest correlation between qPCR gene expression and plasma protein lipocalin 2 (Spearman’s rho = 0.45, P = 0.005, Figure 4b).

Since LCN2 is also a well-described biomarker of acute kidney injury (AKI),(20, 21) we tested the association between expressed LCN2 and ARDS after accounting for creatinine on enrollment. In 49 non-dialysis dependent patients we found that a one-point relative quantity increase in LCN2 was associated with 81% increased odds of ARDS (OR 1.81, 95% CI 1.16-2.81, P = 0.009). This association was unchanged after adjustment for creatinine (OR 1.80, 95% CI 1.16 – 2.80, P=0.009).

Assessment of genes identified in prior studies of sepsis-related ARDS

Using the literature search strategy as described in the methods, we compared our results to those of the two previous whole blood gene expression analysis in ARDS. Three of eight genes reported in the gene signature by Howrylak et al(35) were differentially expressed in the current study with an FDR < 0.25. ADP-ribosylation factor 3 (ARF3), cyclin-dependent kinase inhibitor 1A (CDKN1A) and patatin-like phospholipase domain containing 2 (PNPLA2) were all found to be upregulated in ARDS, although the log2-fold changes for these genes were modest (Table 5). Of the 3 inflammasome genes identified in the study by Dolinay et al,(23) caspase 1 (CASP1) was differentially expressed with an FDR of 0.21 in Model 2. However, while Dolinay reported upregulation of CASP1 in sepsis/ARDS patients compared to those with non-septic systemic inflammatory response syndrome (SIRS), we found downregulation of CASP1 in
ARDS compared with sepsis. In a more recent study using publicly available data from the Howrylak study (35), investigators identified 12 differentially expressed genes. (19) We found that two of the twelve genes were upregulated with FDRs <0.25: histone cluster 1, H4i (HIST1H3H) and CDKN1A (Table 5).

Pathway analysis

Pathway analysis included all differentially expressed genes with FDR < 0.25 and log2 fold change of > 0.26 for model 1. There were 481 genes included, which yielded several key pathogenic pathways shown in Table 6. Pathways that were significantly associated with septic patients with ARDS include genetic programs associated with severe acute respiratory syndrome, infection of the respiratory tract, viral infection and infection of cells. Whereas the respiratory disease pathways include genes that were both up- and down-regulated (as indicated by a non-significant Z-score), the infectious disease pathways contain genes that were markedly upregulated in ARDS. Several apoptosis and cell death annotations were also overrepresented. The pathway results for differentially expressed genes in Model 2 were similar (data not shown).

DISCUSSION

This study compared the early transcriptional differences of whole blood leukocytes in critically ill septic patients with and without ARDS in order to identify key transcriptional changes that might influence the pathogenesis of ARDS. Our study identified several genes and transcriptional pathways that were differentially expressed. Specifically, we identified LCN2, BPI and CD24 and MMP8 as having increased expression in patients with ARDS. Several of these genes have been previously described as pathogenic factors in the development of ARDS, thus supporting this experimental approach to identifying differentially expressed genes. In
addition, we identified several novel genes as well, including OLFM4. The results withstood adjustment for age, sex, batch, the white blood cell count and the presence of direct lung injury. In addition, the same genes were differentially expressed among just the patients with shock, suggesting that the genes identified are not simply markers of concurrent organ failure.

In contrast to previous studies, which compared septic patients with ARDS to healthy controls or other non-septic critically ill patients, our approach may narrow the set of differentially expressed genes and provide economy to identifying novel genes that might regulate the important transition to lung injury in septic patients. Lastly, our results confirmed differential expression of several genes identified in a prior gene expression study of sepsis-related ARDS,(35) providing validity both to the likely relevance of these genes in the pathogenesis of ARDS and to the approach in this study.

The potential mechanisms of the identified genes are summarized in Table 7. Briefly, lipocalin 2 encodes the lipocalin 2 protein, also known as neutrophil gelatinase-associated lipocalin (NGAL), which is expressed by myeloid and epithelial cells. It is a well-described biomarker of acute kidney injury(20, 21) and has known antimicrobial properties due to scavenging of bacterial siderophores and consequent restriction of iron acquisition by gram-negative bacteria.(18, 28, 34, 63) While the role of LCN 2 in ARDS is not fully known, it may mediate the antimicrobial effect of mesenchymal stem cells in experimental gram negative pneumonia.(34) It has also been associated with deactivation of macrophages and induction of an immunosuppressive M2 phenotype in gram-positive pneumonia.(63) The critical role of LCN 2 in the innate immune response and known upregulation in AKI provides biologic plausibility for a role in ARDS pathogenesis. While most early pathogenesis studies have focused on plasma protein biomarkers,(8) the finding of a correlation between lipocalin 2 gene expression and
plasma protein levels demonstrate that, at least in this example, they may be linked. Additionally, we found that the association between expressed LCN2 and ARDS was independent of enrollment creatinine, providing further support that highly expressed LCN2 may play a direct role in ARDS pathogenesis.

The CD24 molecule serves as the granulocyte receptor for platelet P-selectin, a key mediator of deleterious neutrophil-platelet interactions in experimental acute lung injury and may serve as a potential target for platelet depletion or aspirin therapy in patients with ARDS.(24, 47, 74) The BPI protein is stored in neutrophil granules and has a pluripotent antimicrobial effect via neutralization of lipopolysaccharide (LPS) and a highly selective bactericidal effect on gram-negative organisms.(17, 48, 64) In murine LPS acute lung injury models, BPI analog treatment has been associated with a protective effect via decreased pulmonary capillary vessel permeability and reduced pulmonary neutrophil influx suggesting a possible compensatory effect in host response to ARDS.(30)

In addition to confirming genes previously associated with ARDS, the microarray expression approach we utilized provides the opportunity for novel gene discovery. For example, OLFM4 was more highly expressed in patients with sepsis and ARDS. Interestingly, this gene is expressed by neutrophils and may play a critical role in host defense. A recent murine study of OLFM4-/- mice demonstrated enhanced in vivo bacterial clearance and more resistance to bacterial challenge suggesting that OLFM4 may negatively regulate neutrophil bacteriocidal activity in sepsis.(45)

The downregulated genes identified had less consistency across the unadjusted and adjusted models and were not been previously reported in the literature in the pathogenesis of
ARDS. In addition, on qPCR testing of HCAR3 and MME, downregulation was not confirmed. Whether these genes have significance in ARDS pathogenesis remains unclear.

An interesting result of this study was the preponderance of neutrophil-related processes among the identified genes, a finding not explained by circulating neutrophil counts. There was consistency in the candidate genes identified in all unadjusted and adjusted analyses, and the upregulated genes were highly correlated, suggesting that these genes may be co-regulated in the pathogenesis of early ARDS. The premise that neutrophil-related mechanisms are directly involved early in the development of ARDS is supported by both clinical and animal studies. Neutrophils are the first leukocytes to be recruited to sites of inflammation, and although they provide an initial line of defense against invading microorganisms, excessive neutrophil activation can cause collateral damage to host tissues and lead to increased vascular permeability. Although these mechanisms are likely host protective, they may come at the cost of increased capillary-alveolar permeability and impaired of lung function. Thus, the identification of several neutrophil-related genes in this study has biologic plausibility and may provide insights into the early events in ARDS pathogenesis. Further study will be required to understand what involvement these genes may have in the multi-step process of neutrophil recruitment and activation across the pulmonary endothelium, through the lung interstitium and across alveolar epithelium into the air spaces. Considerations for further study should include RNA isolation specifically from circulating neutrophils, and analysis of neutrophil gene expression in bronchoalveolar lavage fluid. Recent advances in microfluidic-based capture of human neutrophils will make these approaches feasible. Circulating white blood cells and absolute neutrophil counts were lower in patients with ARDS, perhaps because of sequestration in the lung microcirculation. Our primary analysis
adjusted for white blood cells to account for the possibility that expression levels of gene would be explained by differences in the number of circulating leukocytes. However, the most differentially expressed genes were the same in unadjusted analysis. Additionally, patients with ARDS had significantly more shock and a higher severity of illness than the septic patients. Therefore, it is not clear whether gene expression differences relate specifically to ARDS in sepsis or whether they represent other mechanisms inherent to a more severely ill patient population.

The finding that three of eight genes in the gene signature proposed by Howrylak and colleagues (35) were differentially expressed in our cohort is notable. While not clearly related to neutrophil biology, these three genes all appear to have relevance in the regulation of cell cycle processes. The finding of genes involved in cell regulation rather than granulocyte activation in the Howrylak study may be explained in part by the later time point in illness of patients in that study, who were enrolled up to 48 hours after admission to the ICU. For instance, it is possible that these genes are involved in later phases of host response in sepsis pathophysiology. (36)

Timing of sample acquisition is an important factor in the evaluation of the pathogenesis of ARDS since ARDS often occurs early in the course of sepsis and the gene expression profile in acute inflammatory diseases have been shown to change rapidly over time. (29, 57)

**Limitations**

This study has several strengths including early enrollment, PCR confirmation, prospective design, detailed phenotyping of patients, and correlation of mRNA expression with plasma protein levels of LCN 2. The study also has some limitations. First, although the false discovery rate is robust for multiple comparisons, our cut-off point of 0.25 suggests that 1 of every 4 genes considered could be a false positive. We adopted this approach with the goal of
identifying candidate genes that will require further validation in a second study population; prior literature suggests that this FDR is reasonable in this context. In order to increase the likelihood that the genes presented were valid, we supplemented our analysis with a literature search on each gene to determine biologic plausibility. It was a reassuring finding that all of the upregulated genes meeting our statistical cut off had been previously described in the literature of ARDS or sepsis, as in the case of OLFM4. The pathway analysis also yielded relevant and plausible results. Second, the sample size of our study is modest, and our data suggest that considerably larger sample sizes would be needed to detect a gene signature of ARDS in early sepsis. However, to our knowledge, this study is the largest gene expression study of sepsis-related ARDS to date. Third, the patients in the ARDS group were more severely ill, at least in part, due to the presence of lung injury. We suspect there is overlap in mechanisms of organ dysfunction in sepsis and though the genes identified may not be exclusive to the development of acute lung injury, the presence of differential expression in patients with shock suggests that severity of systemic illness in sepsis does not explain the genes identified. Lastly, our approach can detect changes only in genes in circulating leukocytes and not in other cells known to be important in the pathogenesis of lung injury such as endothelial and epithelial cells.

Implications and areas for further study

The identification of several neutrophil-related genes identified in the ARDS group is interesting and may represent a novel way to study the role of innate immunity in the development of ARDS. Future approaches using microfluidics for extraction of circulating neutrophils should be investigated to determine how neutrophils are involved in these gene expression changes. In addition, the preponderance of cell death and apoptosis pathways identified is biologically plausible in the pathogenesis of ARDS and may serve as preliminary
Furthermore, collection of RNA in patients both before and after the development of ARDS will be essential for understanding pathologic mechanism at different phases of illness. Finally, heterogeneity in our study population may have attenuated the gene expression differences. Larger studies are warranted and should consider subgroups matched according to severity of illness and shock to further evaluate the possibility that mechanisms of ARDS may differ among patients with variable manifestations of non-pulmonary organ failure. Further study in gene expression programs is likely to increase understanding to subtypes of complex illnesses like ARDS. Recent discoveries in sepsis demonstrate how gene expression subphenotypes can be identified and applied in clinical settings.”

CONCLUSIONS

In summary, the results of this study demonstrate that biologically important mechanisms can be detected utilizing whole blood gene expression in patients early in the course of sepsis, with higher expression of several genes in septic patients with ARDS. These data identify biologically plausible genes with known neutrophil-related functions that are likely to be important in the early pathogenesis of sepsis-related ARDS. In addition, this study validates possible roles for several genes previously identified in microarray studies of sepsis-related ARDS. Future studies focusing on gene expression differences occurring early in the course of sepsis-related ARDS may further delineate neutrophil-related mechanisms involved in the progression to ARDS.
ACKNOWLEDGMENTS:

We thank the patients who participated in the study and their family members. We thank the research staff that assisted with the study. We thank Jason Abbott, in particular, for his laboratory assistance and for measurement of plasma LCN2.

REFERENCES


Table 1. Clinical characteristics in sepsis and ARDS versus sepsis alone.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Sepsis Only</th>
<th>Sepsis + ARDS</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=28</td>
<td>N=29</td>
<td></td>
</tr>
<tr>
<td>Age (mean ± SD)</td>
<td>67 ± 20</td>
<td>59 ± 19</td>
<td>0.11</td>
</tr>
<tr>
<td>Male</td>
<td>16 (57%)</td>
<td>16 (55%)</td>
<td>0.61</td>
</tr>
<tr>
<td>Direct lung injury&lt;sup&gt;b&lt;/sup&gt;</td>
<td>13 (46%)</td>
<td>21 (72%)</td>
<td>0.05</td>
</tr>
<tr>
<td>PAXgene batch 2</td>
<td>17 (61%)</td>
<td>19 (66%)</td>
<td>0.71</td>
</tr>
<tr>
<td>Time to PAXgene draw after ICU admission, mean hours ± SD</td>
<td>9.2 ± 8.1</td>
<td>10.9 ± 9.8</td>
<td>0.46</td>
</tr>
<tr>
<td>WBC, median (IQR) (cells/ul)</td>
<td>14.2 (11.2, 18.6)</td>
<td>11.1 (6.6, 12.6)</td>
<td>0.007</td>
</tr>
<tr>
<td>ANC, median (IQR) (cells/ul)</td>
<td>12.1 (8.5, 16.8)</td>
<td>9.4 (5.0, 11.8)</td>
<td>0.04</td>
</tr>
<tr>
<td>Serum creatinine&lt;sup&gt;c&lt;/sup&gt; mean ± SD (mg/ml)</td>
<td>1.5 ± 1.3</td>
<td>1.7 ± 1.1</td>
<td>0.64</td>
</tr>
<tr>
<td>ESRD</td>
<td>5 (18%)</td>
<td>3 (10%)</td>
<td>0.41</td>
</tr>
<tr>
<td>Cancer&lt;sup&gt;d&lt;/sup&gt;</td>
<td>6 (21%)</td>
<td>15 (52%)</td>
<td>0.02</td>
</tr>
<tr>
<td>APACHE III, mean ± SD</td>
<td>84 ± 31</td>
<td>116 ± 39</td>
<td>0.002</td>
</tr>
<tr>
<td>Shock</td>
<td>12 (43%)</td>
<td>21 (72%)</td>
<td>0.02</td>
</tr>
<tr>
<td>60 day mortality</td>
<td>5 (18%)</td>
<td>9 (31%)</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Data presented as No. (%) unless otherwise indicated. All measures present on enrollment with the exception of mortality.

Definition of abbreviations: ARDS = APACHE = Acute Physiology and Chronic Health Evaluation; Acute Respiratory Distress Syndrome; ESRD = End-stage renal disease; IQR = Interquartile range; SD = Standard deviation; WBC = White blood cell count.

<sup>a</sup> Berlin level of severity defined as mild $\text{PaO}_2/\text{FiO}_2 \geq 200$ to $\leq 300$ mm Hg with PEEP $\geq 5$ cm H$_2$O or continuous positive airway pressure (CPAP) $\geq 5$ cm H$_2$O, moderate $\text{PaO}_2/\text{FiO}_2 100$ to $\leq 200$ mm Hg with PEEP $\geq 5$ cm H$_2$O, and severe $\text{PaO}_2/\text{FiO}_2 \leq 100$ mm Hg with PEEP $\geq 5$ cm H$_2$O.(51)

<sup>b</sup> Direct lung injury is defined as ARDS risk factor of pneumonia or aspiration

<sup>c</sup> In patients without end-stage renal disease

<sup>d</sup> Includes, solid metastatic, solid non-metastatic, leukemia, lymphoma and multiple myeloma

<sup>e</sup> In survivors
Table 2. Top differentially expressed genes in sepsis and ARDS versus sepsis alone in unadjusted analysis

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene Description</th>
<th>ENTREZ ID</th>
<th>Log₂ Fold change</th>
<th>Raw P-value</th>
<th>FDR</th>
<th>No. Pubmed Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLFM4*</td>
<td>Olfactomedin 4</td>
<td>10562</td>
<td>1.495</td>
<td>0.002</td>
<td>0.21</td>
<td>0</td>
</tr>
<tr>
<td>CD24*</td>
<td>CD24 molecule</td>
<td>100133941</td>
<td>1.152</td>
<td>0.004</td>
<td>0.24</td>
<td>7</td>
</tr>
<tr>
<td>LCN2*</td>
<td>Lipocalin 2</td>
<td>3934</td>
<td>1.141</td>
<td>0.003</td>
<td>0.21</td>
<td>41</td>
</tr>
<tr>
<td>BPI*</td>
<td>Bactericidal/permeability-increasing protein</td>
<td>671</td>
<td>1.071</td>
<td>0.001</td>
<td>0.16</td>
<td>27</td>
</tr>
<tr>
<td>RBP7</td>
<td>Retinol binding protein 7, cellular</td>
<td>116362</td>
<td>-1.007</td>
<td>0.002</td>
<td>0.07</td>
<td>0</td>
</tr>
<tr>
<td>UTS2</td>
<td>Urotensin 2</td>
<td>10911</td>
<td>-1.032</td>
<td>0.004</td>
<td>0.24</td>
<td>0</td>
</tr>
</tbody>
</table>

* Gene with neutrophil-related mechanism.

PubMed Search completed March 12, 2014
Table 3. Top differentially expressed genes in sepsis and ARDS versus sepsis alone in *Model 1* adjusted analysis

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene Description</th>
<th>ENTREZ ID</th>
<th>Log₂ Fold change</th>
<th>Raw P-value</th>
<th>FDR</th>
<th>No. Pubmed Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upregulated in ARDS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLFM4*</td>
<td>Olfactomedin 4</td>
<td>10562</td>
<td>1.482</td>
<td>0.004</td>
<td>0.17</td>
<td>0</td>
</tr>
<tr>
<td>LCN2*</td>
<td>Lipocalin 2</td>
<td>3934</td>
<td>1.245</td>
<td>0.002</td>
<td>0.13</td>
<td>41</td>
</tr>
<tr>
<td>CD24*</td>
<td>CD24 molecule</td>
<td>100133941</td>
<td>1.207</td>
<td>0.005</td>
<td>0.18</td>
<td>7</td>
</tr>
<tr>
<td>BPI*</td>
<td>Bactericidal/permeability-increasing protein</td>
<td>671</td>
<td>1.113</td>
<td>0.001</td>
<td>0.11</td>
<td>27</td>
</tr>
<tr>
<td><strong>Downregulated in ARDS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCAR3</td>
<td>Hydroxycarboxylic acid receptor 3</td>
<td>8843</td>
<td>-1.018</td>
<td>0.002</td>
<td>0.14</td>
<td>0</td>
</tr>
<tr>
<td>MME</td>
<td>Membrane metallo-endopeptidase</td>
<td>4311</td>
<td>-1.230</td>
<td>0.007</td>
<td>0.21</td>
<td>0</td>
</tr>
</tbody>
</table>

Adjusted for age, sex, batch and white blood cell count

* Gene with neutrophil-related mechanism.

* PubMed Search completed March 12, 2014
Table 4. Top differentially expressed genes in sepsis and ARDS versus sepsis alone in Model 2 adjusted analysis.

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene Description</th>
<th>ENTREZ ID</th>
<th>Log2 Fold change</th>
<th>Raw P-value</th>
<th>FDR</th>
<th>No. Pubmed Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upregulated in ARDS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMP 8*</td>
<td>Matrix metallopeptidase 8</td>
<td>4317</td>
<td>1.642</td>
<td>0.007</td>
<td>0.19</td>
<td>31</td>
</tr>
<tr>
<td>OLFM4*</td>
<td>Olfactomedin 4</td>
<td>10562</td>
<td>1.596</td>
<td>0.003</td>
<td>0.14</td>
<td>0</td>
</tr>
<tr>
<td>LCN2*</td>
<td>Lipocalin 2</td>
<td>3934</td>
<td>1.385</td>
<td>0.001</td>
<td>0.10</td>
<td>41</td>
</tr>
<tr>
<td>CD24*</td>
<td>CD24 molecule</td>
<td>100133941</td>
<td>1.370</td>
<td>0.002</td>
<td>0.12</td>
<td>7</td>
</tr>
<tr>
<td>HP</td>
<td>Haptoglobin</td>
<td>3240</td>
<td>1.275</td>
<td>0.001</td>
<td>0.10</td>
<td>90</td>
</tr>
<tr>
<td>BPI*</td>
<td>Bactericidal/permeability-increasing protein</td>
<td>671</td>
<td>1.214</td>
<td>0.001</td>
<td>0.09</td>
<td>27</td>
</tr>
<tr>
<td>RETN</td>
<td>Resistin</td>
<td>56729</td>
<td>1.112</td>
<td>&lt;0.001</td>
<td>0.08</td>
<td>0b</td>
</tr>
<tr>
<td>TCN1</td>
<td>Transcobalamin 1</td>
<td>6947</td>
<td>1.039</td>
<td>0.002</td>
<td>0.13</td>
<td>0</td>
</tr>
<tr>
<td><strong>Downregulated in ARDS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNORD1A</td>
<td>Small nucleolar RNA, C/D box 1A</td>
<td>677848</td>
<td>-1.048</td>
<td>0.007</td>
<td>0.19</td>
<td>0</td>
</tr>
<tr>
<td>HCAR3</td>
<td>Hydroxycarboxylic acid receptor 3</td>
<td>8843</td>
<td>-1.049</td>
<td>0.003</td>
<td>0.13</td>
<td>0</td>
</tr>
<tr>
<td>LOC642373</td>
<td>Contactin associated protein-like 3B pseudogene</td>
<td>642373</td>
<td>-1.119</td>
<td>0.007</td>
<td>0.19</td>
<td>0</td>
</tr>
<tr>
<td>RBP7</td>
<td>Retinol binding protein 7, cellular</td>
<td>116362</td>
<td>-1.122</td>
<td>&lt;0.001</td>
<td>0.04</td>
<td>0</td>
</tr>
<tr>
<td>SNORD64</td>
<td>Small nucleolar RNA, C/D box 64</td>
<td>347686</td>
<td>-1.177</td>
<td>0.003</td>
<td>0.14</td>
<td>0</td>
</tr>
<tr>
<td>MME</td>
<td>Membrane metallo-endopeptidase</td>
<td>4311</td>
<td>-1.595</td>
<td>&lt;0.001</td>
<td>0.08</td>
<td>0</td>
</tr>
</tbody>
</table>

Adjusted for age, sex, batch, WBC and direct lung injury risk factor
Grey rows denote new genes not identified in the unadjusted or Model 1 adjusted analyses.

*a PubMed search completed March 12, 2014
b 4 references identified on search. None relevant to RETN gene and ARDS or sepsis
* Gene with neutrophil-related mechanism.
Table 5. Association with candidate genes differentially expressed in prior gene expression studies of sepsis-related ARDS*

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene Description</th>
<th>ENTREZ ID</th>
<th>Model 1 Log2 Fold change</th>
<th>FDR</th>
<th>Model 2 Log2 Fold change</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Howrylak et al(35)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTH1</td>
<td>ferritin, heavy polypeptide 1</td>
<td>2495</td>
<td>NA*</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARF3</td>
<td>ADP-ribosylation factor 3</td>
<td>377</td>
<td>0.222</td>
<td>0.248</td>
<td>0.255</td>
<td>0.17</td>
</tr>
<tr>
<td>BTG2</td>
<td>BTG family, member 2</td>
<td>7832</td>
<td>-0.075</td>
<td>0.80</td>
<td>-0.083</td>
<td>0.78</td>
</tr>
<tr>
<td>NQO2</td>
<td>NAD(P)H dehydrogenase, quinone 2</td>
<td>4835</td>
<td>-0.148</td>
<td>0.71</td>
<td>-0.167</td>
<td>0.68</td>
</tr>
<tr>
<td>CDKN1A</td>
<td>cyclin-dependent kinase inhibitor 1A</td>
<td>1026</td>
<td>0.160</td>
<td>0.20</td>
<td>0.157</td>
<td>0.22</td>
</tr>
<tr>
<td>PNPLA2</td>
<td>patatin-like phospholipase domain containing 2</td>
<td>57104</td>
<td>0.247</td>
<td>0.22</td>
<td>0.246</td>
<td>0.23</td>
</tr>
<tr>
<td>NPEPL1</td>
<td>aminopeptidase-like 1</td>
<td>79716</td>
<td>0.028</td>
<td>0.89</td>
<td>0.033</td>
<td>0.87</td>
</tr>
<tr>
<td>CREBZF</td>
<td>CREB/ATF bZIP transcription factor</td>
<td>58487</td>
<td>-0.153</td>
<td>0.57</td>
<td>-0.179</td>
<td>0.50</td>
</tr>
<tr>
<td>Dolinay et al(23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CASP1</td>
<td>caspase 1, apoptosis-related cysteine peptidase</td>
<td>834</td>
<td>-0.374</td>
<td>0.21</td>
<td>-0.331</td>
<td>0.28</td>
</tr>
<tr>
<td>IL18</td>
<td>interleukin 18</td>
<td>3606</td>
<td>0.071</td>
<td>0.92</td>
<td>0.090</td>
<td>0.90</td>
</tr>
<tr>
<td>IL1B</td>
<td>interleukin 1, beta</td>
<td>3553</td>
<td>-0.363</td>
<td>0.51</td>
<td>-0.251</td>
<td>0.69</td>
</tr>
<tr>
<td>Chen et al(19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOPX</td>
<td>HOP homeobox</td>
<td>84525</td>
<td>0.020</td>
<td>0.93</td>
<td>0.017</td>
<td>0.95</td>
</tr>
<tr>
<td>CYBRD1</td>
<td>Cytochrome b reductase 1</td>
<td>79901</td>
<td>0.255</td>
<td>0.57</td>
<td>0.209</td>
<td>0.67</td>
</tr>
<tr>
<td>UPB1</td>
<td>Ureidopropionase, beta</td>
<td>51733</td>
<td>0.068</td>
<td>0.71</td>
<td>0.087</td>
<td>0.61</td>
</tr>
<tr>
<td>OCLN</td>
<td>Occludin</td>
<td>100506</td>
<td>-0.266</td>
<td>0.68</td>
<td>-0.354</td>
<td>0.55</td>
</tr>
<tr>
<td>C21orf7</td>
<td>Chromosome 21 open reading frame 7</td>
<td>59611</td>
<td>-0.037</td>
<td>0.89</td>
<td>-0.033</td>
<td>0.91</td>
</tr>
<tr>
<td>HIST2H4B</td>
<td>Histone cluster 1, H4h</td>
<td>8365</td>
<td>-0.066</td>
<td>0.90</td>
<td>-0.064</td>
<td>0.91</td>
</tr>
<tr>
<td>TREM1</td>
<td>Triggering receptor expressed on myeloid cells 1</td>
<td>54210</td>
<td>-0.510</td>
<td>0.36</td>
<td>-0.580</td>
<td>0.29</td>
</tr>
<tr>
<td>HIST1H3H</td>
<td>Histone cluster 1, H3i</td>
<td>8354</td>
<td>0.551</td>
<td>0.24</td>
<td>0.525</td>
<td>0.28</td>
</tr>
<tr>
<td>CDKN1A</td>
<td>cyclin-dependent kinase inhibitor 1A</td>
<td>1026</td>
<td>0.16</td>
<td>0.20</td>
<td>0.157</td>
<td>0.22</td>
</tr>
<tr>
<td>BTNL8</td>
<td>Butyrophilin-like 8</td>
<td>79908</td>
<td>-0.116</td>
<td>0.78</td>
<td>-0.089</td>
<td>0.85</td>
</tr>
<tr>
<td>HLA-DQB1</td>
<td>Major histocompatibility complex class II, DQ beta 1</td>
<td>3119</td>
<td>-0.140</td>
<td>0.84</td>
<td>-0.197</td>
<td>0.76</td>
</tr>
<tr>
<td>CDKN1C</td>
<td>cyclin-dependent kinase inhibitor 1C</td>
<td>1028</td>
<td>0.068</td>
<td>0.50</td>
<td>0.060</td>
<td>0.57</td>
</tr>
</tbody>
</table>

Genes shaded in grey are differentially expressed in current study with FDR < 0.25

* Not included in the Affymetrix transcript cluster ID annotation “v1.na33.1.hg19”
Table 6. Top 15 disease annotations overrepresented in IPA pathway analysis for Model 1 includes 481 genes differentially expressed with FDR <0.25 and log2 fold change > 0.2625

<table>
<thead>
<tr>
<th>Categories</th>
<th>Diseases or Functions</th>
<th>Annotation</th>
<th>P-Value</th>
<th>z-score</th>
<th># genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell Death and Survival</td>
<td>cell death</td>
<td></td>
<td>1.48E-15</td>
<td>-0.640</td>
<td>174</td>
</tr>
<tr>
<td>Cell Cycle, Cellular Movement</td>
<td>cytokinesis</td>
<td></td>
<td>2.42E-14</td>
<td>1.048</td>
<td>25</td>
</tr>
<tr>
<td>Cell Death and Survival</td>
<td>apoptosis</td>
<td></td>
<td>3.87E-14</td>
<td>-0.491</td>
<td>144</td>
</tr>
<tr>
<td>Cell Cycle</td>
<td>M phase</td>
<td></td>
<td>1.15E-13</td>
<td>1.048</td>
<td>29</td>
</tr>
<tr>
<td>Infectious Disease, Respiratory Disease</td>
<td>severe acute respiratory syndrome</td>
<td></td>
<td>5.25E-12</td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>Infectious Disease, Respiratory Disease</td>
<td>infection of respiratory tract</td>
<td></td>
<td>8.38E-12</td>
<td></td>
<td>22</td>
</tr>
<tr>
<td>Cancer</td>
<td>malignant neoplasm of abdomen</td>
<td></td>
<td>1.17E-11</td>
<td>1.037</td>
<td>199</td>
</tr>
<tr>
<td>Cell Death and Survival</td>
<td>necrosis</td>
<td></td>
<td>4.97E-11</td>
<td>-0.738</td>
<td>132</td>
</tr>
<tr>
<td>Cellular Growth and Proliferation</td>
<td>proliferation of cells</td>
<td></td>
<td>5.46E-11</td>
<td>2.091</td>
<td>169</td>
</tr>
<tr>
<td>Cancer, Organ Development, Organismal Injury and Abnormalities, Renal and Urological Disease</td>
<td>renal cancer</td>
<td></td>
<td>2.05E-10</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>Infectious Disease</td>
<td>Viral Infection</td>
<td></td>
<td>2.51E-10</td>
<td>3.544</td>
<td>95</td>
</tr>
<tr>
<td>Cancer</td>
<td>Cancer</td>
<td></td>
<td>6.13E-10</td>
<td>1.294</td>
<td>281</td>
</tr>
<tr>
<td>Cancer, Organ Development, Organismal Injury and Abnormalities, Renal and Urological Disease</td>
<td>renal-cell carcinoma</td>
<td></td>
<td>6.99E-10</td>
<td></td>
<td>26</td>
</tr>
<tr>
<td>Hematological Disease</td>
<td>anemia</td>
<td></td>
<td>1.93E-09</td>
<td>-2.022</td>
<td>31</td>
</tr>
<tr>
<td>Cell Death and Survival</td>
<td>cell death of tumor cell lines</td>
<td></td>
<td>1.20E-08</td>
<td>-0.582</td>
<td>82</td>
</tr>
</tbody>
</table>

Disease annotations shaded in grey relevant to sepsis and infection-related respiratory complications
**Table 7.** Potential functions with relevance to ARDS among differentially expressed genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Potential Mechanism</th>
</tr>
</thead>
</table>
| LCN2 | • Expressed by myeloid and epithelial cells  
      • Antimicrobial properties via scavenging of bacterial siderophores with restriction of iron acquisition by gram negative bacteria (18, 28, 34)  
      • Macrophage deactivation with induction of immunosuppressive M2 phenotype (63)  
      • Biomarker of acute kidney injury (20, 21) |
| CD24 | • Encodes CD 24 sialoglycoprotein expressed on mature granulocytes and B cells  
      • Granulocyte receptor for platelet P-selectin, a mediator of neutrophil-platelet interactions in ARDS (24, 47, 74) |
| BPI  | • Stored in neutrophil granules with pluripotent antimicrobial effect via neutralization of lipopolysaccharide  
      • Selective bactericidal effect on gram-negative organisms (17, 48, 64)  
      • BPI analog associated with decreased pulmonary capillary vessel permeability and reduced pulmonary neutrophil influx (30) |
| MMP 8| • Inactive protein stored within neutrophil granules activated by autolytic cleavage resulting in collagen degradation (4)  
      • Imbalance between MMP-8 and its natural tissue inhibitor implicated in ARDS pathogenesis (27, 38, 40, 53) |
| HP   | • Acute phase protein modulate immune response via binding of free hemoglobin (9)  
      • Disruption of iron metabolism and effect oxidative stress (32, 73) |
**FIGURE LEGENDS**

**Figure 1.** Flow diagram of inclusion and exclusion criteria.

**Figure 2.** Scatterplot matrix demonstrating Spearman correlations between differentially expressed genes in sepsis-related ARDS. Figure 2a includes differentially expressed genes in Model 1. Figure 2b presents the 8 additional genes identified in Model 2.

**Figure 3.** Normalized results of confirmation quantitative PCR on log2 scale comparing gene expression in patients with sepsis and ARDS versus sepsis alone. Y-axis are relative quantities (RQ) of the message levels relative to the lowest value, set arbitrarily at RQ=1.

**Figure 4.** Scatterplot matrix demonstrating the correlations between (4a) normalized microarray gene expression levels of lipocalin 2 and measured plasma protein levels of LCN 2 (Spearman’s rho is 0.67, P < 0.0001) and between (4b) normalized qPCR expression levels of lipocalin 2 and measured plasma protein levels of lipocalin 2 (Spearman’s rho is 0.45, P = 0.005).
This diagram illustrates the total number of patients enrolled in the ongoing EARLI study during the selected enrollment period and the reasons for excluding patients based on predetermined criteria. Following this process, there were 29 patients with sepsis and ARDS and 28 sepsis patients without ARDS.

*Abbreviations:* EARLI=Early assessment of renal and lung injury; ARDS=Acute Respiratory distress syndrome; CXR=chest radiograph
## FIGURE 2

Figure 2a

<table>
<thead>
<tr>
<th></th>
<th>0.91</th>
<th>0.79</th>
<th>0.81</th>
<th>-0.39&lt;sup&gt;a&lt;/sup&gt;</th>
<th>-0.59</th>
<th>0.00&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLFM4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LCN2</td>
<td>0.83</td>
<td>0.81</td>
<td>-0.44</td>
<td>-0.67</td>
<td>-0.05&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>D24</td>
<td>0.82</td>
<td></td>
<td>-0.49</td>
<td>-0.53</td>
<td>-0.04&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>BPI</td>
<td>-0.31&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.48</td>
<td></td>
<td>-0.08&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCAR3</td>
<td>0.50</td>
<td>0.41&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MME</td>
<td>0.14&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

P-value < 0.001 unless otherwise stated

<sup>a</sup> P-value ≥ 0.001 but < 0.05
<sup>b</sup> P-value ≥ 0.05 but < 0.25
<sup>c</sup> P-value ≥ 0.25
P-value < 0.001 unless otherwise stated

\( ^a \) P-value ≥ 0.001 but < 0.05

\( ^b \) P-value ≥ 0.05 but < 0.25

\( ^c \) P-value ≥ 0.25
** means $P < 0.01$ *** means $P < 0.001$
FIGURE 4

Figure 4a