

RESEARCH

Open Access

Host lung gene expression patterns predict infectious etiology in a mouse model of pneumonia

Scott E Evans*^{1,2}, Michael J Tuvim^{1,2}, Jiexin Zhang³, Derek T Larson¹, Cesar D García⁴, Sylvia Martinez Pro⁴, Kevin R Coombes³ and Burton F Dickey^{1,2}

Abstract

Background: Lower respiratory tract infections continue to exact unacceptable worldwide mortality, often because the infecting pathogen cannot be identified. The respiratory epithelia provide protection from pneumonias through organism-specific generation of antimicrobial products, offering potential insight into the identity of infecting pathogens. This study assesses the capacity of the host gene expression response to infection to predict the presence and identity of lower respiratory pathogens without reliance on culture data.

Methods: Mice were inhalationally challenged with *S. pneumoniae*, *P. aeruginosa*, *A. fumigatus* or saline prior to whole genome gene expression microarray analysis of their pulmonary parenchyma. Characteristic gene expression patterns for each condition were identified, allowing the derivation of prediction rules for each pathogen. After confirming the predictive capacity of gene expression data in blinded challenges, a computerized algorithm was devised to predict the infectious conditions of subsequent subjects.

Results: We observed robust, pathogen-specific gene expression patterns as early as 2 h after infection. Use of an algorithmic decision tree revealed 94.4% diagnostic accuracy when discerning the presence of bacterial infection. The model subsequently differentiated between bacterial pathogens with 71.4% accuracy and between non-bacterial conditions with 70.0% accuracy, both far exceeding the expected diagnostic yield of standard culture-based bronchoscopy with bronchoalveolar lavage.

Conclusions: These data substantiate the specificity of the pulmonary innate immune response and support the feasibility of a gene expression-based clinical tool for pneumonia diagnosis.

Background

Pneumonias result in substantial mortality, causing more premature death and disability worldwide than any other disease [1]. Unfortunately, while patient survival depends upon the rapid identification of infecting pathogens [2], the means for prompt and accurate diagnoses of pulmonary infections remain inadequate.

Despite widespread acceptance as the diagnostic tool of choice for unexplained pulmonary infiltrates [3-5], fiberoptic bronchoscopy with bronchoalveolar lavage (BAL) provides an unambiguous diagnosis in only 25-

51% of cases [2,4,6-9]. The diagnostic utility of BAL is predicated on culturing pathogens from lavage effluent, without accounting for ongoing antibiotic therapy, non-pathogenic microbial colonization, or the technical challenge of navigating the bronchoscope into involved airways. Molecular techniques, such as antigen detection and polymerase chain reaction (PCR) testing, enhance BAL sensitivity for a subset of pathogens, but still often fail to explain infiltrates [7].

Often regarded as passive gas exchange barriers, the active responses of the lungs are critical to protection from infections. In the presence of inflammatory stimuli, the respiratory epithelia rapidly recruit inflammatory cells and undergo remarkable structural and functional

* Correspondence: seevans@mdanderson.org

¹ Department of Pulmonary Medicine, University of Texas - M. D. Anderson Cancer Center, Houston, Texas, USA

Full list of author information is available at the end of the article

changes [10-13], including the release of pathogen-specific antimicrobial products [14-16].

Even in the absence of an adaptive immune system, lower metazoans like *Drosophila melanogaster* selectively respond to different classes of microorganisms following pathogen detection with conserved pattern recognition receptors [17,18]. Similarly, stereotyped pathogen-specific host innate immune responses are also observed from human dendritic cells [19], human monocytic cells [20-24], human endothelial cells [25], murine microglial cells [26], and murine jejunal epithelial cells [27]. Based upon these multiply observed tailored responses and the inflammatory capacity of pulmonary epithelium [12,28], we hypothesized that the lungs also respond selectively to different pathogens. In order to pursue the potential to achieve superior diagnostic utility in a timely manner, we interrogated this selective response to determine the etiology of pneumonias without reliance on culture data.

Methods

Animals and reagents

Unless otherwise specified, reagents were obtained from Sigma (St Louis, MO). All experiments were approved by the M. D. Anderson Cancer Center Institutional Animal Care and Use Committee. Specific pathogen free BALB/c mice were purchased from Harlan (Indianapolis, IN) and used in experiments at five to eight weeks old.

Infection Model

To achieve simultaneous exposure of large numbers of mice to respiratory pathogens, mice were placed in a nebulization chamber that was sealed except for an efflux limb that vented to a low resistance filter in a biohazard hood. An AeroMist CA-209 compressed gas nebulizer (CIS-US, Inc., Bedford, MA) was used to aerosolize pathogen suspensions, driven by 10 L/min of room air and supplemented with 5% CO₂ to promote maximal ventilation and homogeneous exposure throughout the lungs, as we have previously described [29-32]. While it is conceivable that exposure of mice to increased inspired CO₂ concentrations might alter gene expression, our experience supports prior reports that this promotes pathogen deposition in the lungs [33,34], and our strategy involves differential gene expression analysis where all mice are exposed to the same CO₂ environment, thus no differential effects should be detected.

Organisms

For bacterial pathogens, the inocula were targeted to an LD₇₅ by 48 h after infection. After growth to log phase, *Streptococcus pneumoniae* serotype 4, and *Pseudomonas aeruginosa* strain PA103 were each suspended in phosphate buffered saline (PBS) and delivered by aerosol. A standardized nebulization of 10 ml pathogen suspension

over one hour to achieve the desired lethality required concentrations of approximately 1 × 10¹⁰ CFU/per ml *S. pneumoniae* and approximately 1 × 10¹¹ CFU/ml of *P. aeruginosa*, as we have previously described [29,31].

Because *Aspergillus fumigatus* is not lethal in non-immunosuppressed BALB/c mice, we delivered the maximal reproducible concentration of organisms as limited by viscosity. This dose was 1 × 10⁹ conidia/ml, as determined using a standard hemacytometer. Conidia of strain Af293 were stored as frozen stock (1 × 10⁹ conidia/ml) in 20% glycerol in PBS. One ml of stock was plated on yeast extract agar plates at 37°C in 5% CO₂ for 3 days, then harvested by gentle scraping in PBS containing 0.1% Tween-20, and the suspension was filtered through 40 μm filters, centrifuged at 2,500 × g for 10 min, washed, resuspended in 10 ml PBS and aerosolized over 60 min, identical to the bacterial infections. To confirm both pulmonary deposition and infective capacity of the pathogen, additional mice were challenged with the same *A. fumigatus* protocol with or without prior cyclophosphamide and cortisol immunosuppression, as previously described [31].

A sham intervention group was treated with 10 ml PBS nebulized over 60 min under the same conditions used for infectious challenges.

Cytokine response to pathogen challenge

At designated time points after infection, mice were anesthetized and their tracheas were exposed. BAL was performed and lavage effluent cytokine concentrations were determined by ELISA, as described [29,30].

Gene expression analysis

At designated time points after infection, gene expression microarray analysis was performed on lung homogenates from mice after challenge following leukoreduction by repeated BAL and vascular perfusion with sterile PBS [31,32]. Lungs were excised and homogenized, total RNA was extracted, and amplified cRNA was hybridized to Illumina Sentrix Mouse-6 BeadChips (Illumina, Inc., San Diego, CA). All primary data were deposited at the NCBI Gene Expression Omnibus <http://www.ncbi.nlm.nih.gov/geo/>, accession GSE15869) consistent with MIAME standards (see Additional File 1).

Blinded challenges

To test the predictive ability of the gene expression data, three blinded investigators (SEE, MJT, BFD) were independently challenged to identify the infectious conditions based on gene expression patterns without reliance on culture data. After identifying characteristic changes for each condition in the gene expression analysis, the investigators were provided the data from only six transcripts that were each believed to be uniquely altered by one of the potential infectious conditions. In order to identify

potentially discriminating transcripts and to assign cutoff values for a diagnostic panel we used two approaches. First, after confirming that there was no overlap of signal intensity between 2 standard deviations below a differentially upregulated gene and 2 standard deviations above the next highest condition for that transcript, we assigned a cutoff value for a positive test at 1 standard deviation below the mean signal intensity for the transcript in question. As a second approach, we created receiver operating characteristic (ROC) curves for each potentially discriminating transcript, selected from the list of differentially expressed genes. Two potentially predictive genes for each of the three infections were hand-selected for the panel, and the investigators were instructed to predict the pathogen based on the pre-stated rules (Additional File 2). Investigators were instructed to infer that a sample was from the sham group if the values did not meet criteria for one of the infections.

Computer Algorithm

A computer algorithm was devised to automate the prediction of infecting organisms, based on the 18 h microarray data described above. The predictive model is a decision tree, with the first branch a decision between lungs infected with a bacterial pathogen and those not infected with bacteria. The sequential decisions are between *S. pneumoniae* and *P. aeruginosa* in the bacterial branch and between *A. fumigatus* and sham in the non-bacterial branch. Transcripts with predictive power to discern between branches were identified by fitting a linear model for each transcript, then the infectious condition of each blinded sample was sequentially predicted based on the expression of 1 to 21 discrete transcripts, with each transcript "voting" for one side of the decision tree (e.g., predicting either "bacterial" or "not bacterial"). To avoid ties when using majority vote rule, only odd numbers of predictor genes were allowed (see Additional File 1).

Results

Infectious pneumonia model

Consistent with our prior observations [29-31], our bacterial pneumonia model yielded highly reproducible mortality (Figure 1A). No mortality was observed following fungal challenges or sham treatment. We confirmed delivery of infective conidia through the observation of highly reproducible mortality at the same inoculum for immunosuppressed mice (Figure 1B), and serial dilution culture of lung homogenates showed deposition of approximately 3×10^6 conidia per *A. fumigatus*-challenged mouse.

Proteomic comparison

We initially suspected that lung cytokine responses to different pathogens might be diagnostically predictive. To

test this, we compared the BAL concentration of 16 inflammatory cytokines by ELISA (Additional File 3) to determine whether this approach would allow discernment of the conditions. Representative examples of IFN- γ , TNF- α , IL-6, and CCL-17 are shown in Figure 2 to be strongly induced by *P. aeruginosa* infection, with lesser induction by other infections. No cytokines were uniquely induced by any other pathogen.

Infection-induced gene expression changes

Since the protein-level cytokine response only differentiated *P. aeruginosa* from the other conditions, we interrogated the transcriptional response of differently infected lungs. Gene expression differences emerged very early after infection. Using an extremely rigorous false discovery rate (FDR) $< 1 \times 10^{-7}$, we identified 20 differentially expressed genes (DEGs) at our earliest investigated time point, 2 h after challenge. By 6 h after challenge, this number had increased to 4,274 DEGs, nearly 10% of the 45,992 oligonucleotides probed. By unsupervised clustering, the samples tended to assemble themselves into condition-specific groups even at this early time point, with grossly recognizable patterns already developing (Figure 3B).

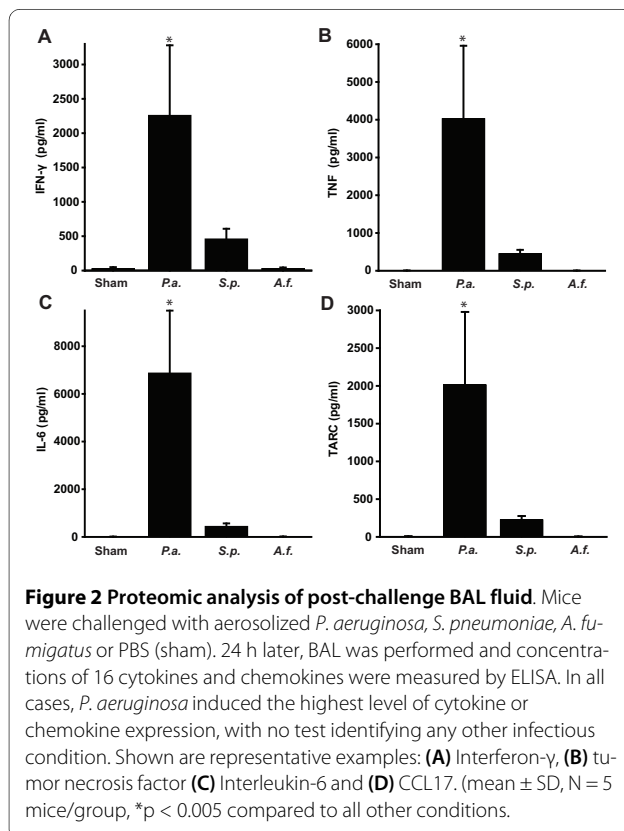
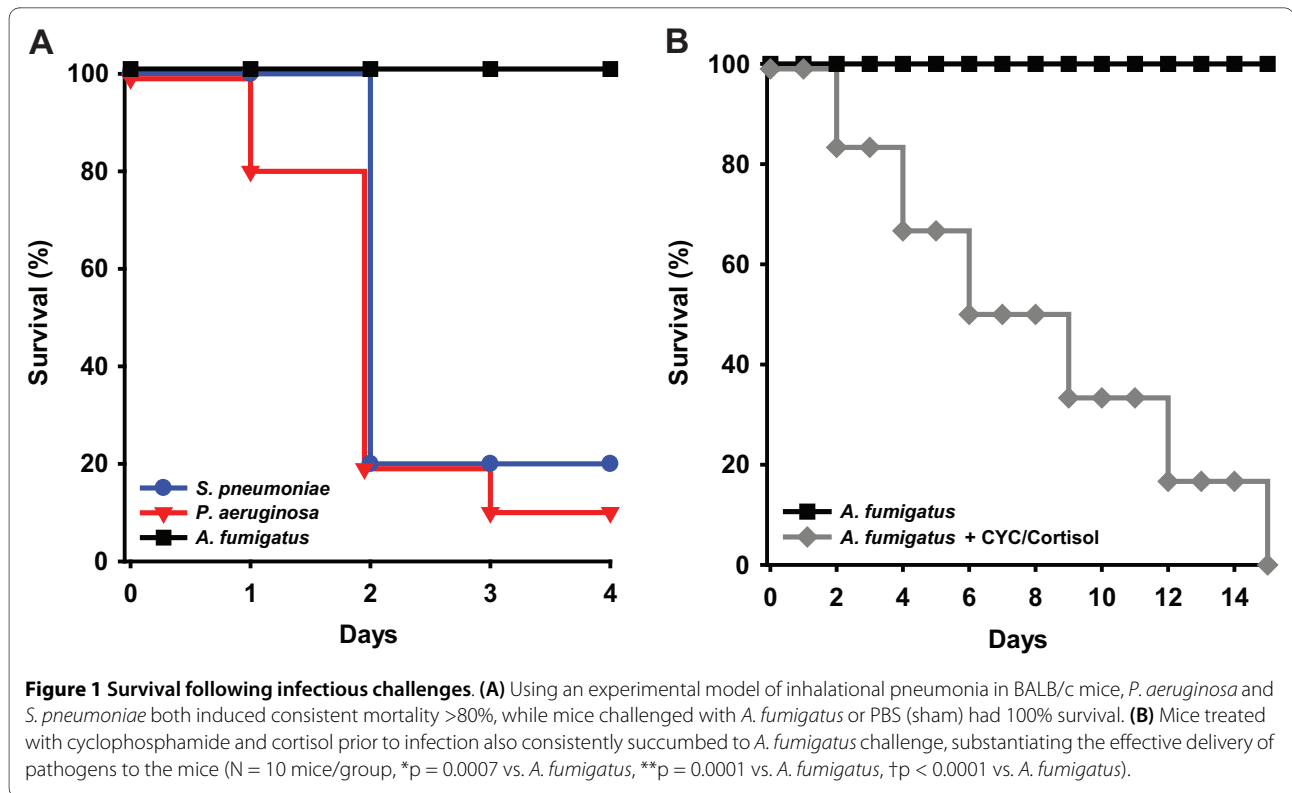
Over the course of 12 to 18 h, the total number of DEGs at FDR of 1×10^{-7} decreased to 367, but even greater condition-specific clustering was observed than at 6 h. Of these 367 DEGs, 179 were differentially expressed at both 6 h and 18 h time points. Notably, while the total number of DEGs decreased over time, the average fold-change of the remaining DEGs was generally increased.

Figure 3B demonstrates the temporal effect on gene expression in this model. The 30 most strongly differentially expressed genes at 18 h were analyzed at earlier time points, revealing progressive intensification of the gene expression patterns. Of these 30 DEGs, 18 were also differentially expressed at 6 h.

Condition-specific transcripts

By 18 h after challenge, unsupervised clustering resulted in all of the specimens correctly segregating themselves by pathogen (Figure 4A). After identifying patterns associated with each infectious condition, we focused on individual transcripts with each condition. The 367 DEGs at 18 h were sorted according to pathogen specificity (Figure 4B). Not surprisingly, the two conditions that caused mortality induced more gene expression changes than did *A. fumigatus*. However, each condition induced unique changes, and by lessening the FDR requirements, these numbers further increase.

Manual review of the 367 DEGs identified unique transcript changes for each pathogen that were included in a predictive panel. Strategies using either the magnitude of differential expression or ROC curve performance were equally efficacious for defining the prediction rule cut-off

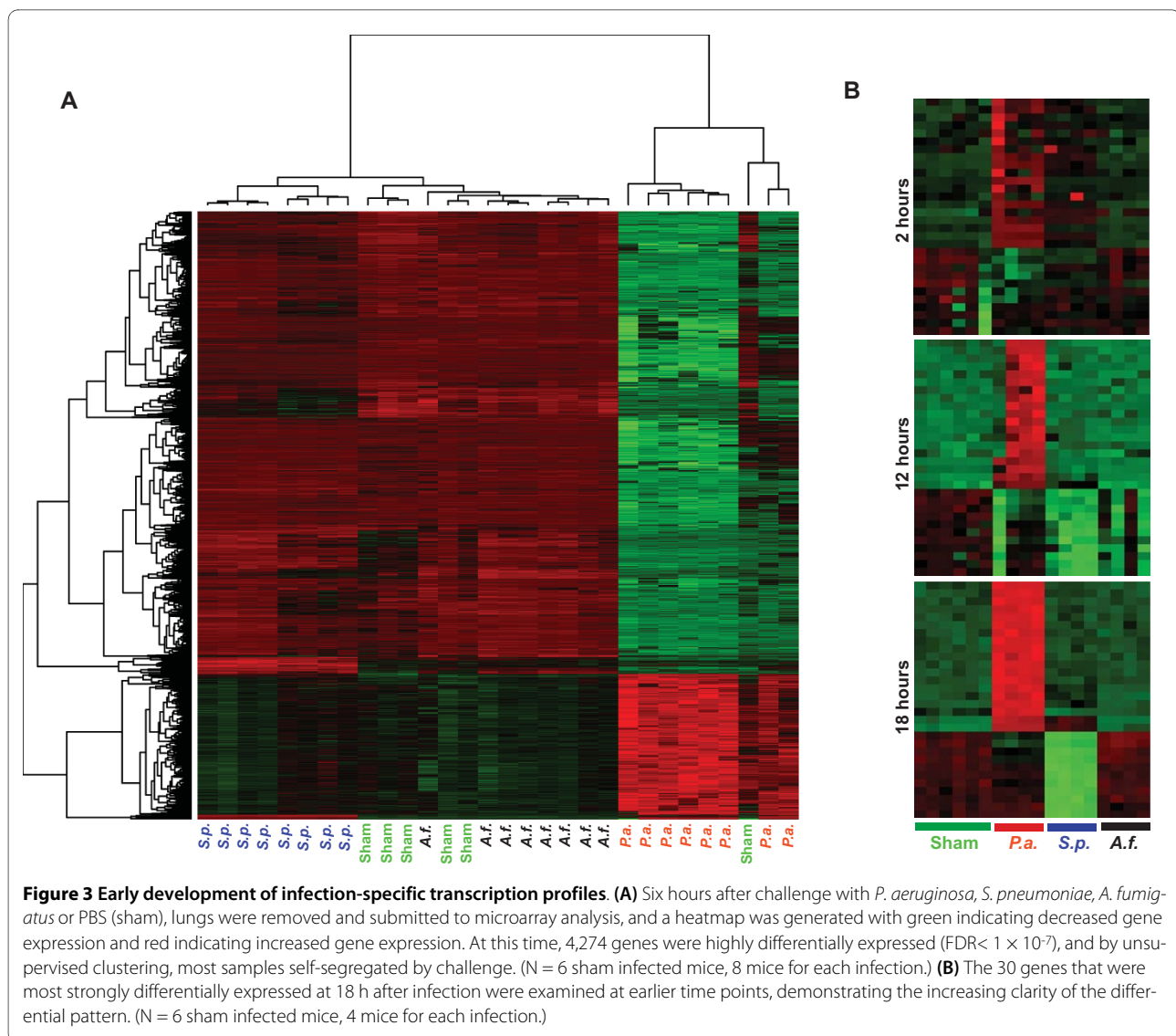


values. As shown in Additional File 4, each included transcript yielded a cutoff that achieved 100% sensitivity and 100% specificity in the 18 h training set (i.e., area under the ROC curve = 1.0). Additional Files 2 and 5 show the panel of transcripts, the prediction rules, the data provided to the blinded investigators, and their predictions. Blinded review of 18 samples at 18 h after infection resulted in 100% correct categorization of infectious conditions for all three reviewers.

When we applied these prediction rules to 18 unique samples from a validation dataset, however, the prediction accuracy dropped to only 44.4%. As shown in Additional File 6, Additional File 7 and Additional File 8, there was congruity of the blinded investigators insofar as samples were most often either correctly predicted by all three investigators or incorrectly predicted by all three investigators. No statistically significant patterns emerged among the incorrect predictions.

Computerized prediction of infectious conditions

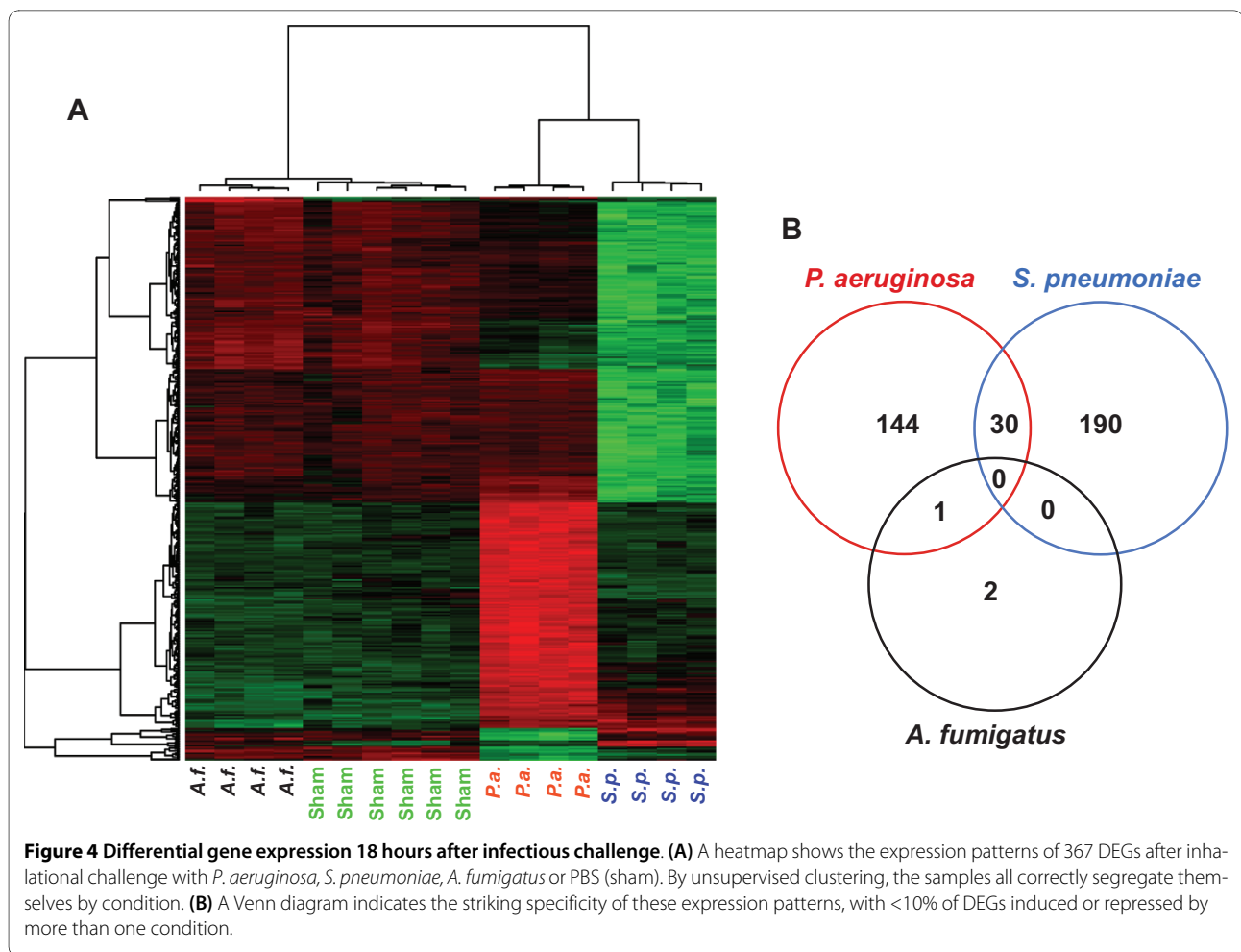
Since a small panel of hand selected transcripts predicted infectious conditions as well or better than traditional cultures historically perform, we sought to automate the process of prediction. We devised a multiply branching decision tree algorithm that first separated bacterial infections (*S. pneumoniae* and *P. aeruginosa*) from non-bacterial conditions (*A. fumigatus* and sham). We identified 4,799 transcripts from the training set that could dis-



tinguish these two groups. Using our predetermined criteria for predictor transcripts, we found that *Ccl4* (chemokine C-C motif ligand 4) performed most robustly, correctly classifying all training set samples as bacteria or non-bacteria. We also found individual transcripts with very high predictive accuracy for subsequent branches of the decision tree. *Ccl3* (chemokine C-C motif ligand 3) expression always separated *S. pneumoniae* infection from *P. aeruginosa* in the training set. A single gene, *Ttn* (titin), discriminated between *A. fumigatus* and sham in 90% of the samples, reflecting all but one sample accurately categorized by the transcript. Notably, the sham sample that was inaccurately categorized as *A. fumigatus* by *Ttn* was also predicted to be *A. fumigatus* using multiple other transcripts, and inspection of the overall gene expression profile appeared more consistent with *A. fumigatus* than sham. This raises the possibility

that the mouse was inadvertently or incidentally infected with fungus. If true, the *Ttn*-based categorization would be 100% correct for this branch point, as well.

After identifying the most discriminant transcripts from the training set, we tested the 18 h predictor genes at other time points. Gene expression data from lungs 2 h, 6 h, and 12 h after infection were combined into a single group, then infectious predictions were made according to the algorithm. As described in Additional File 1, we used increasing odd numbers of "voting" transcripts up to 21. The prediction accuracy for discriminating bacterial infections from non-bacterial conditions was 78% for 2 h specimens, 100% for 6 h specimens, and 89% for 12 h specimens using 15 transcripts as predictors (Figure 5A). Performance of the model appeared to stabilize around 15 "voting" transcripts per branch point, with only a modest increase to 83% accuracy at 2 h using 21 predic-

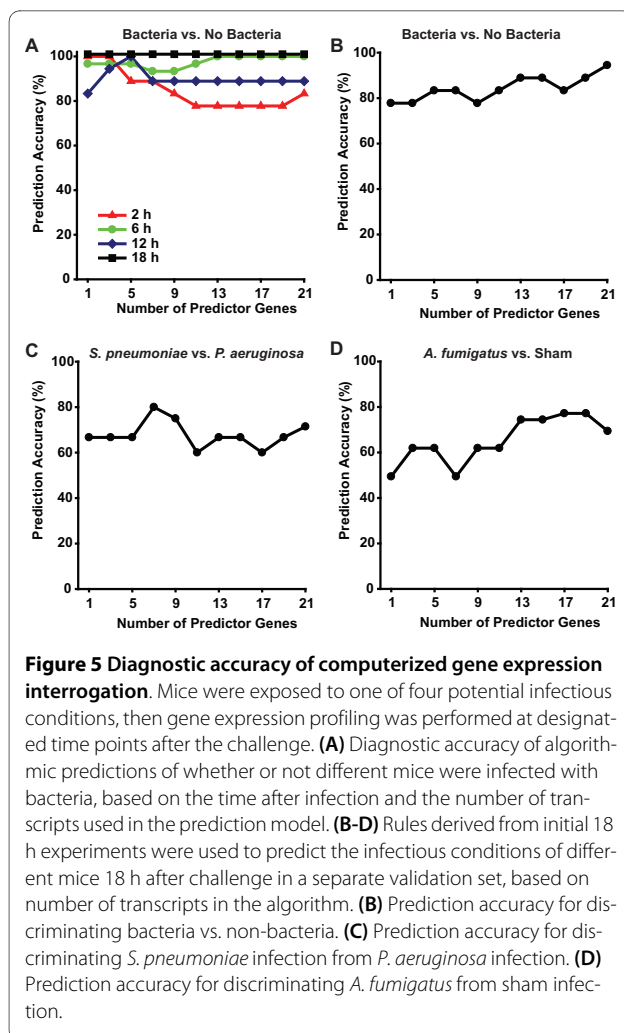


tors. There were no changes in the other time points with more than 15 predictors. The accuracy of diagnoses predicted by the second branch point was 50%, 30%, 55% and 95% for the 2, 6, 12, and 18 h specimens using 15 predictors, respectively.

The decision tree model was then tested against a unique (validation) set of gene expression data from lung homogenates collected 18 h after challenge. Using the same algorithm, the correct prediction of bacterial vs. non-bacterial status was made with 89% accuracy with 15 predictor genes and with 94.4% accuracy with 21 predictors (Figure 5B). Discrimination of *S. pneumoniae* vs. *P. aeruginosa* and of *A. fumigatus* vs. sham was achieved with >70% accuracy (Figure 5C and 5D). We again found that increasing the number of "voting" transcripts improved accuracy, with stabilization around 15 transcripts. The effect of adding additional predictor transcripts was minimal for separating the bacterial conditions from each other, but increasing from 3 to 15 transcripts correctly reclassified several samples from *A. fumigatus* to sham.

Discussion

The informative value of host responses is increasingly recognized to differentiate between clinically confounding conditions [35]. Markers of generic inflammation have been used for decades to hint at the presence of inflammatory and infectious diseases [36,37]. More recently, host response elements have been studied to aid identification of life-threatening diseases, such as sTREM and procalcitonin in respiratory infections and sepsis [38-41]. Efforts are underway to characterize pulmonary conditions as diverse as interstitial lung diseases, pulmonary vascular diseases and asthma based on gene expression analysis [42-46]. Diagnostic host responses to *Mycobacterium tuberculosis*, are increasingly described [47-49]. Differential gene expression has been reported in the lungs following different infections [50] and gene expression profiling of leukocytes has been proposed to provide prognostic insights in the setting of lung infection [51]. However, to the best of our knowledge, this report is the first to describe a means of identifying etiologic agents of infectious pneumonia based solely on the host gene expression response.



Because of the potential ease of sampling and abundance, we first sought to discriminate between infectious conditions based on BAL cytokine levels. Using a panel of 16 cytokines, *P. aeruginosa*-infected mice were consistently differentiated from the other three conditions. This is consistent with the recent report of McConnell and colleagues who found that a panel of 18 cytokines could discriminate *P. aeruginosa*-from *S. pneumoniae*-infected mice [52]. However, while we identified a robust cytokine signature for one pathogen, we were unable to discern between the non-pseudomonal conditions by that method. Further proteomic analysis for non-cytokine host response elements may discriminate between the conditions, but our prior experience resolving low abundance peptides from BAL fluid [29] suggests that the technical challenges would offset the enhanced diagnostic capacity. Therefore, we elected to investigate host response specificity using gene expression analysis.

Our gene expression data suggest that host responses are sufficiently specific to discriminate between conditions that may be indistinguishable, such as different

infectious pneumonias. While there appears to be a modest early peak of non-specific inflammation, we were surprised to identify such efficient discrimination by as early as 2 h after challenge. By 6 h after challenge, there was a robust response that waned in number of DEGs by 12 h, but clearly increased in signal amplitude of the persisting transcript changes. This durable signal increased to the 18 h time point and allowed for consistent blinded diagnoses. Remarkably, fewer than 10% of the 367 DEGs at 18 h were induced by more than one infectious condition (none by all three). Further, we found no evidence that the different infections simply induced the same gene expression patterns at different paces, rather each condition resulted in a unique gene expression profile. These findings attest to the high specificity of the host response. While the number of *Aspergillus*-regulated transcripts was low compared to the bacteria-induced DEGs, these findings are consistent with the finding of DeGregorio, et al. [53], and of Huang, et al. [19], when investigating fungus-induced gene expression in *Drosophila* and in human dendritic cells, respectively. Based on these results, we hypothesize that human lung gene expression patterns on clinical biopsy specimens will demonstrate similar specificity.

In order to systematize the otherwise subjective process of pattern-identification and to automate the process for efficiency, we devised a computerized algorithm to test whether gene expression data could predict subjects' infectious states. From a practical perspective, this strategy allowed simultaneous assessment of massive numbers of transcript permutations. More importantly, it provided diagnostic accuracy far better than that typically encountered clinically with traditional culture-based diagnostic strategies, and outperformed diagnostic predictions based on gene expression of hand selected transcripts.

The algorithm was intentionally structured as a decision tree. This allows for determination of the most relevant questions first, for sequentially increasing refinement of answers, and for the flexibility to add new branch points. In this case, based differences in available treatment options, we felt the most clinically important issue was to differentiate subjects with bacterial pneumonia from those without bacterial pneumonia. The program provided great accuracy in answering this question. The model was also robust for the secondary questions, though less so.

Typical of preliminary investigations, these data have limitations to their generalizability. Comparison of three organisms from different pathologic classes makes it impossible to know whether the effects observed are species-specific or broader effects of the group. This will be assessed in future comparisons to other members of the same classes. By design, the decision tree algorithm

allows for exactly this type of modification. It is also possible that some of the gene expression changes observed in the *A. fumigatus*-infected animals may represent the effects of their immunosuppression. Given the clinical focus of our cancer center, future studies of potential drug effects will be a high priority.

Another advantage of interrogating gene expression profiles in suspected pneumonia is that it allows somewhat compartmentalized analyses of different cellular elements of the host response. Because the lungs were leukoreduced by bronchoalveolar lavage and vascular perfusion, the data presented here largely reflect responses of the epithelium. Expression patterns from simultaneously harvested alveolar macrophages will be separately analyzed and presented. Although the cellular purity is incomplete, this approach may be viewed as a preliminary model of the clinical situation where RNA can be separately obtained from epithelial cells by brushing and from alveolar macrophages by BAL. Such discrete analyses may be applied to identifying etiologies of other pulmonary condition, as well.

It could be argued that the samples were harvested sooner after initiation of infection than would be clinically possible. However, our model causes diffuse and uniform infection of the lungs, whereas clinical pneumonia generally begins with a localized infection that progresses spatially and temporally. Therefore, a clinical specimen harvested from the most recently involved lung segments will also be newly infected. Further, our observations of increasing signal intensity over time suggest that a durable diagnostic pattern will be identifiable at later stages. This will require confirmation in future, longer term studies.

Conclusions

The early and accurate diagnosis of the etiology of pneumonia would be of great clinical benefit. These findings suggest that it may be feasible to harness the host response to inform clinicians of a patient's infectious state when pneumonia is suspected. We anticipate that this will allow for development of a clinically-relevant tool, as well as providing new insights into differences between normal and ineffective host responses to infections.

Additional material

Additional file 1 Supplemental Methods. Text file containing additional experimental and data handling methods details.

Additional file 2 Supplemental Table 2. Prediction rules for manually selected transcripts. Table of rules provided to blinded investigators for predicting infectious challenges.

Additional file 3 Supplemental Table 1. BAL fluid cytokine levels 24 h after infection with different pathogens. Table of BAL cytokine levels for each mouse following pathogen challenges.

Additional file 4 Supplemental Figure 1. Individual transcripts discriminate between infectious conditions. Receiver operating characteristic (ROC) curves for transcripts from the 18 hour post-infection training set that discriminate *P. aeruginosa*, *S. pneumoniae*, and *A. fumigatus* from the other three potential conditions.

Additional file 5 Supplemental Table 3. Training set data provided to blinded investigators. Table of gene expression data from the training set provided to blinded investigators.

Additional file 6 Supplemental Table 4. Validation set data provided to blinded investigators. Table of the gene expression data from the validation set provided to blinded investigators.

Additional file 7 Supplemental Table 5. Condition predictions from validation set. Table of the predictions made by each blinded investigator for each subject.

Additional file 8 Supplemental Table 6. Validation set performance of predictions based on hand-selected transcripts. Table of blinded prediction performance.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SEE participated in design, performance and analysis of the infectious and microarray experiments, and wrote the manuscript. MJT participated in design, performance and analysis of the infectious experiments. JZ participated in the analysis of microarray data and writing of the manuscript. DTL participated in the performance of infectious experiments and performance microarray experiments. CDG participated analysis of microarray data. SMP participated in performance of infectious experiments. KRC participated in the analysis of microarray data and writing of the manuscript. BFD participated in design and analysis of the infectious experiments and writing of the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

The authors wish to thank Dr. Molly Bray, Baylor College of Medicine, for her assistance with the performance of the microarray studies. This work was supported by the National Institutes of Health [KL2 RR02419] and institutional funds from the University of Texas M. D. Anderson Cancer Center to Dr Evans. Bioinformatics resources for this work was supported by a Cancer Center Support Grant [P30 CA016672] to the University of Texas M. D. Anderson Cancer Center. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Details

¹Department of Pulmonary Medicine, University of Texas - M. D. Anderson Cancer Center, Houston, Texas, USA, ²Center for Lung Inflammation and Infection, Texas A&M Institute for Biosciences and Technology, Houston, Texas, USA, ³Department of Bioinformatics and Computational Biology, University of Texas - M. D. Anderson Cancer Center, Houston, Texas, USA and ⁴Tecnológico de Monterrey School of Medicine, Monterrey, Nuevo León, Mexico

Received: 8 April 2010 Accepted: 23 July 2010

Published: 23 July 2010

References

1. WHO: The World Health Report 2004 -- Changing History. Geneva: World Health Organization; 2004.
2. Rano A, Agusti C, Jimenez P, Angrill J, Benito N, Danes C, Gonzalez J, Rovira M, Pumarola T, Moreno A, *et al.*: Pulmonary infiltrates in non-HIV immunocompromised patients: a diagnostic approach using non-invasive and bronchoscopic procedures. *Thorax* 2001, **56**(5):379-387.
3. Shorr AF, Susla GM, O'Grady NP: Pulmonary infiltrates in the non-HIV-infected immunocompromised patient: etiologies, diagnostic strategies, and outcomes. *Chest* 2004, **125**(1):260-271.
4. White P, Bonacum JT, Miller CB: Utility of fiberoptic bronchoscopy in bone marrow transplant patients. *Bone Marrow Transplant* 1997, **20**(8):681-687.
5. Bissinger AL, Einsele H, Hamprecht K, Schumacher U, Kandolf R, Loeffler J, Aepinus C, Bock T, Jahn G, Hebart H: Infectious pulmonary complications after stem cell transplantation or chemotherapy:

- diagnostic yield of bronchoalveolar lavage. *Diagn Microbiol Infect Dis* 2005, **52**(4):275-280.
6. Hohenthal U, Kiworr M, Genitsariotis R, Zeidler D, Lorenz J: **Role of bronchoalveolar lavage in immunocompromised patients with pneumonia treated with a broad spectrum antibiotic and antifungal regimen.** *Thorax* 2001, **56**(2):115-120.
 7. Hohenthal U, Itala M, Salonen J, Sipilä J, Rantakokko-Jalava K, Meurman O, Nikoskelainen J, Vainionpää R, Kotilainen P: **Bronchoalveolar lavage in immunocompromised patients with haematological malignancy--value of new microbiological methods.** *Eur J Haematol* 2005, **74**(3):203-211.
 8. Kasow KA, King E, Rochester R, Tong X, Srivastava DK, Horwitz EM, Leung W, Woodard P, Handgretinger R, Hale GA: **Diagnostic yield of bronchoalveolar lavage is low in allogeneic hematopoietic stem cell recipients receiving immunosuppressive therapy or with acute graft-versus-host disease: the St. Jude experience, 1990-2002.** *Biol Blood Marrow Transplant* 2007, **13**(7):831-837.
 9. Shannon VR, Andersson BS, Lei X, Champlin RE, Kontoyiannis DP: **Utility of early versus late fiberoptic bronchoscopy in the evaluation of new pulmonary infiltrates following hematopoietic stem cell transplantation.** *Bone Marrow Transplant* 2009, **45**(4):647-655.
 10. Evans CM, Williams OW, Tuvim MJ, Nigam R, Mixides GP, Blackburn MR, DeMayo FJ, Burns AR, Smith C, Reynolds SD, *et al.*: **Mucin is produced by clara cells in the proximal airways of antigen-challenged mice.** *Am J Respir Cell Mol Biol* 2004, **31**(4):382-394.
 11. Williams OW, Sharafkhaneh A, Kim V, Dickey BF, Evans CM: **Airway mucus: From production to secretion.** *Am J Respir Cell Mol Biol* 2006, **34**(5):527-536.
 12. Evans SE, Hahn PY, McCann F, Kottom TJ, Pavlovic ZV, Limper AH: **Pneumocystis cell wall beta-glucans stimulate alveolar epithelial cell chemokine generation through nuclear factor-kappaB-dependent mechanisms.** *Am J Respir Cell Mol Biol* 2005, **32**(6):490-497.
 13. Hippenstiel S, Opitz B, Schmeck B, Suttrop N: **Lung epithelium as a sentinel and effector system in pneumonia--molecular mechanisms of pathogen recognition and signal transduction.** *Respir Res* 2006, **7**:97.
 14. Rogan MP, Geraghty P, Greene CM, O'Neill SJ, Taggart CC, McElvaney NG: **Antimicrobial proteins and polypeptides in pulmonary innate defence.** *Respir Res* 2006, **7**:29.
 15. Bals R, Hiemstra PS: **Innate immunity in the lung: how epithelial cells fight against respiratory pathogens.** *Eur Respir J* 2004, **23**(2):327-333.
 16. Hickman-Davis JM, Fang FC, Nathan C, Shepherd VL, Voelker DR, Wright JR: **Lung surfactant and reactive oxygen-nitrogen species: antimicrobial activity and host-pathogen interactions.** *Am J Physiol Lung Cell Mol Physiol* 2001, **281**(3):L517-523.
 17. Lemaitre B, Reichhart JM, Hoffmann JA: **Drosophila host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms.** *Proc Natl Acad Sci USA* 1997, **94**(26):14614-14619.
 18. Hoffmann JA: **The immune response of Drosophila.** *Nature* 2003, **426**(6962):33-38.
 19. Huang Q, Liu D, Majewski P, Schulte LC, Korn JM, Young RA, Lander ES, Hacohen N: **The plasticity of dendritic cell responses to pathogens and their components.** *Science* 2001, **294**(5543):870-875.
 20. Boldrick JC, Alizadeh AA, Diehn M, Dudoit S, Liu CL, Belcher CE, Botstein D, Staudt LM, Brown PO, Relman DA: **Stereotyped and specific gene expression programs in human innate immune responses to bacteria.** *Proc Natl Acad Sci USA* 2002, **99**(2):972-977.
 21. Nau GJ, Richmond JF, Schlesinger A, Jennings EG, Lander ES, Young RA: **Human macrophage activation programs induced by bacterial pathogens.** *Proc Natl Acad Sci USA* 2002, **99**(3):1503-1508.
 22. McCaffrey RL, Fawcett P, O'Riordan M, Lee KD, Havell EA, Brown PO, Portnoy DA: **A specific gene expression program triggered by Gram-positive bacteria in the cytosol.** *Proc Natl Acad Sci USA* 2004, **101**(31):11386-11391.
 23. Chaussabel D, Allman W, Mejias A, Chung W, Bennett L, Ramilo O, Pascual V, Palucka AK, Banchereau J: **Analysis of significance patterns identifies ubiquitous and disease-specific gene-expression signatures in patient peripheral blood leukocytes.** *Ann N Y Acad Sci* 2005, **1062**:146-154.
 24. Ramilo O, Allman W, Chung W, Mejias A, Ardura M, Glaser C, Wittkowski KM, Piqueras B, Banchereau J, Palucka AK, *et al.*: **Gene expression patterns in blood leukocytes discriminate patients with acute infections.** *Blood* 2007, **109**(5):2066-2077.
 25. Matussek A, Strindhall J, Stark L, Rohde M, Geffers R, Buer J, Kihlstrom E, Lindgren PE, Lofgren S: **Infection of human endothelial cells with Staphylococcus aureus induces transcription of genes encoding an innate immunity response.** *Scand J Immunol* 2005, **61**(6):536-544.
 26. McKimmie CS, Roy D, Forster T, Fazakerley JK: **Innate immune response gene expression profiles of N9 microglia are pathogen-type specific.** *J Neuroimmunol* 2006, **175**(1-2):128-141.
 27. Knight PA, Pemberton AD, Robertson KA, Roy DJ, Wright SH, Miller HR: **Expression profiling reveals novel innate and inflammatory responses in the jejunal epithelial compartment during infection with Trichinella spiralis.** *Infect Immun* 2004, **72**(10):6076-6086.
 28. Wright TW, Johnston CJ, Harmsen AG, Finkelstein JN: **Analysis of cytokine mRNA profiles in the lungs of Pneumocystis carinii-infected mice.** *Am J Respir Cell Mol Biol* 1997, **17**(4):491-500.
 29. Clement CG, Evans SE, Evans CM, Hawke D, Kobayashi R, Reynolds PR, Moghaddam SJ, Scott BL, Melicoff E, Adachi R, *et al.*: **Stimulation of lung innate immunity protects against lethal pneumococcal pneumonia in mice.** *Am J Respir Crit Care Med* 2008, **177**(12):1322-1330.
 30. Clement CG, Tuvim MJ, Evans CM, Tuvim DM, Dickey BF, Evans SE: **Allergic lung inflammation alters neither susceptibility to Streptococcus pneumoniae infection nor inducibility of innate resistance in mice.** *Respir Res* 2009, **10**:70.
 31. Evans SE, Scott BL, Clement CG, Pawlik J, Bowden MG, Hook M, Kontoyiannis DP, Lewis RE, LaSala PR, Peterson JW, *et al.*: **Stimulation of lung innate immunity protects mice broadly against bacterial and fungal pneumonia.** *Am J Respir Cell Molec Biol* 2010, **42**:40-50.
 32. Tuvim MJ, Evans SE, Clement CG, Dickey BF, Gilbert BE: **Augmented lung inflammation protects against influenza A pneumonia.** *PLoS ONE* 2009, **4**(1):e4176.
 33. Koshkina NV, Knight V, Gilbert BE, Golunski E, Roberts L, Waldrep JC: **Improved respiratory delivery of the anticancer drugs, camptothecin and paclitaxel, with 5% CO2-enriched air: pharmacokinetic studies.** *Cancer Chemother Pharmacol* 2001, **47**(5):451-456.
 34. Newton PE, Pfledderer C: **Measurement of the deposition and clearance of inhaled radiolabeled particles from rat lungs.** *J Appl Toxicol* 1986, **6**(2):113-119.
 35. Jenner RG, Young RA: **Insights into host responses against pathogens from transcriptional profiling.** *Nat Rev Microbiol* 2005, **3**(4):281-294.
 36. Tibble JA, Bjarnason I: **Non-invasive investigation of inflammatory bowel disease.** *World J Gastroenterol* 2001, **7**(4):460-465.
 37. Jaye DL, Waites KB: **Clinical applications of C-reactive protein in pediatrics.** *Pediatr Infect Dis J* 1997, **16**(8):735-746. quiz 746-737
 38. Chao WC, Wang CH, Chan MC, Chow KC, Hsu JY, Wu CL: **Predictive value of serial measurements of sTREM-1 in the treatment response of patients with community-acquired pneumonia.** *J Formos Med Assoc* 2007, **106**(3):187-195.
 39. Gibot S, Cravoisy A, Levy B, Bene MC, Faure G, Bollaert PE: **Soluble triggering receptor expressed on myeloid cells and the diagnosis of pneumonia.** *N Engl J Med* 2004, **350**(5):451-458.
 40. Gibot S, Kolopp-Sarda MN, Bene MC, Bollaert PE, Lozniewski A, Mory F, Levy B, Faure GC: **A soluble form of the triggering receptor expressed on myeloid cells-1 modulates the inflammatory response in murine sepsis.** *J Exp Med* 2004, **200**(11):1419-1426.
 41. Reinhart K, Brunkhorst FM: **Meta-analysis of procalcitonin for sepsis detection.** *Lancet Infect Dis* 2007, **7**(8):500-502. author reply 502-503
 42. Brass DM, Tomfohr J, Yang IV, Schwartz DA: **Using mouse genomics to understand idiopathic interstitial fibrosis.** *Proc Am Thorac Soc* 2007, **4**(1):92-100.
 43. Bull TM, Coldren CD, Geraci MW, Voelkel NF: **Gene expression profiling in pulmonary hypertension.** *Proc Am Thorac Soc* 2007, **4**(1):117-120.
 44. Hansel NN, Diette GB: **Gene expression profiling in human asthma.** *Proc Am Thorac Soc* 2007, **4**(1):32-36.
 45. Lewis CC, Yang JY, Huang X, Banerjee SK, Blackburn MR, Baluk P, McDonald DM, Blackwell TS, Nagabhushanam V, Peters W, *et al.*: **Disease-specific gene expression profiling in multiple models of lung disease.** *Am J Respir Crit Care Med* 2008, **177**(4):376-387.
 46. Pennings JL, Kimman TG, Janssen R: **Identification of a common gene expression response in different lung inflammatory diseases in rodents and macaques.** *PLoS One* 2008, **3**(7):e2596.
 47. Keller C, Lauber J, Blumenthal A, Buer J, Ehlers S: **Resistance and susceptibility to tuberculosis analysed at the transcriptome level:**

- lessons from mouse macrophages. *Tuberculosis (Edinb)* 2004, **84**(3-4):144-158.
48. McGarvey JA, Wagner D, Bermudez LE: **Differential gene expression in mononuclear phagocytes infected with pathogenic and non-pathogenic mycobacteria.** *Clin Exp Immunol* 2004, **136**(3):490-500.
49. Wang JP, Rought SE, Corbeil J, Guiney DG: **Gene expression profiling detects patterns of human macrophage responses following *Mycobacterium tuberculosis* infection.** *FEMS Immunol Med Microbiol* 2003, **39**(2):163-172.
50. Rosseau S, Hocke A, Mollenkopf H, Schmeck B, Suttorp N, Kaufmann SH, Zerrahn J: **Comparative transcriptional profiling of the lung reveals shared and distinct features of *Streptococcus pneumoniae* and influenza A virus infection.** *Immunology* 2007, **120**(3):380-391.
51. McDunn JE, Husain KD, Polpitiya AD, Burykin A, Ruan J, Li Q, Schierding W, Lin N, Dixon D, Zhang W, *et al.*: **Plasticity of the systemic inflammatory response to acute infection during critical illness: development of the riboleukogram.** *PLoS One* 2008, **3**(2):e1564.
52. McConnell KW, McDunn JE, Clark AT, Dunne WM, Dixon DJ, Turnbull IR, Dipasco PJ, Osberghaus WF, Sherman B, Martin JR, *et al.*: ***Streptococcus pneumoniae* and *Pseudomonas aeruginosa* pneumonia induce distinct host responses.** *Crit Care Med* 2010, **38**(1):223-241.
53. De Gregorio E, Spellman PT, Rubin GM, Lemaitre B: **Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays.** *Proc Natl Acad Sci USA* 2001, **98**(22):12590-12595.

doi: 10.1186/1465-9921-11-101

Cite this article as: Evans *et al.*, Host lung gene expression patterns predict infectious etiology in a mouse model of pneumonia *Respiratory Research* 2010, **11**:101

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

