Expression Profiling in Granulomatous Lung Disease

Edward S. Chen and David R. Moller

Division of Pulmonary and Critical Care Medicine, Department of Medicine, The Johns Hopkins University, Baltimore, Maryland

Granulomatous lung diseases, such as sarcoidosis, hypersensitivity pneumonitis, Wegener's granulomatosis, and chronic beryllium disease, along with granulomatous diseases of known infectious etiologies, such as tuberculosis, are major causes of morbidity and mortality throughout the world. Clinical manifestations of these diseases are highly heterogeneous, and the determinants of disease susceptibility and clinical course (e.g., resolution vs. chronic, progressive fibrosis) are largely unknown. The underlying pathogenic mechanisms of these diseases also remain poorly understood. Within this context, these diseases have been approached using genomic and proteomic technologies to allow us to identify patterns of gene/protein expression that track with clinical disease or to identify new pathways involved in disease pathogenesis. The results from these initial studies highlight the potential for these "omics" approaches to reveal novel insights into the pathogenesis of granulomatous lung disease and provide new tools to improve diagnosis, clinical classification, course prediction, and response to therapy. Realizing this potential will require collaboration among multidisciplinary groups with expertise in the respective technologies, bioinformatics, and clinical medicine for these complex diseases.

Keywords: genomics; granuloma; proteomics; sarcoidosis; tuberculosis

The application of expression profiling to granulomatous lung disease is in its infancy. Sarcoidosis, hypersensitivity pneumonitis (HP), Wegener's granulomatosis (WG), and chronic beryllium-induced lung disease (CBD) have been the targets of a limited number of genomic and proteomic studies. In contrast, mycobacterial disease represents an infectious granulomatous lung disease that has been the subject of more intense application of genomic and proteomic approaches applied to mycobacterial organisms and host responses to these microbial agents (1). Genomic, proteomic, lipidomic, and metabolomic profiling offer complementary information, and new approaches for handling the scope of data generated by the respective technologies are being developed (2, 3).

EXPRESSION PROFILING IN SPECIFIC GRANULOMATOUS LUNG DISEASES

Sarcoidosis

Evidence from multiple studies suggests that granulomatous inflammation in sarcoidosis is associated with dominant T helper 1 (Th1) cytokine expression involving IFN- γ , interleukin (IL)-2, and the Th1 immunoregulatory cytokines IL-12 and IL-18, at least early in the disease course (4). Tumor necrosis factor (TNF)- α and a variety of other cytokines and chemokines that are involved in experimental granulomatous inflammation have been shown to be differentially expressed in sarcoidosis. These

initial studies relied on an analysis of the expression of individual or small sets of genes or proteins using specific gene-targeted assays such as Northern blot, polymerase chain reaction (PCR), *in situ* hybridization, Western blot, immunohistochemistry, or ELISA tests (5). However, even when considered together, these studies have not synthesized a comprehensive picture of the orchestration of granulomatous inflammation in sarcoidosis by inflammatory mediators, nor have they provided specific expression patterns that are useful in clinical phenotyping or course prediction.

A more comprehensive approach to disease-specific expression profiling was possible with the development of mRNA differential display. For example, Gaede and colleagues used differential display to determine patterns of gene expression in leukocytes isolated from bronchoalveolar lavage (BAL) cells obtained from five patients with sarcoidosis, five patients with tuberculosis, and five control individuals (6). From 42 different primer combinations, gel electrophoresis revealed 2,498 transcripts detected by PCR, of which 93.5% were common to all patient groups, and far fewer PCR transcripts were specifically associated with sarcoidosis (3.0%), tuberculosis (1.7%), or both (1.8%). In a related study, gene expression induced by Kveim-Siltzbach test reagent in BAL cells from two patients with sarcoidosis was studied (7). Differential display with 11 primer combinations revealed 598 PCR products, of which 1.3% were specifically associated with incubation with Kveim-Siltzbach test reagent. These two studies provided evidence that the proportion of genes specifically up-regulated in sarcoidosis is less than a few percent of the total number of gene products that were examined.

A limited mRNA differential display study by Wiwien and coworkers involved BAL cells obtained from 18 patients with sarcoidosis versus eight patients with other lung diseases (HP, tuberculosis, idiopathic pulmonary fibrosis (IPF), pneumoconiosis, lupus, and bronchiolitis obliterans with organizing pneumonia) (8). Using 19 sets of primers, they identified 12 PCR products that were consistently identified from a majority of these patients. Of these, three PCR products were detected most consistently among patients with sarcoidosis; two of these were sequenced and found to be CD44 and TNF- α , and the third was determined to be a novel gene, the identity of which has not been reported.

Newer microarray gene expression techniques allow for the aggregate analysis of total steady-state gene expression and have been used in a few studies of sarcoidosis (9). Rutherford and coworkers performed microarray analysis of peripheral blood mononuclear cell (PBMC) and found that gene expression in acute-onset pulmonary sarcoidosis was associated with the differential expression of cell growth factors, including heparin-binding epidermal growth factor-like growth factor, endothelial cell growth factor 1, platelet-derived endothelial cell growth factor, c-sis platelet-derived growth factor 2, and vascular endothelial growth factor (10). They also found several TNF- α -related (TNF- α , TNF receptor 1, A20, and TNF-stimulated gene-14[TSG-14]) and other cell-death-related genes (RAIDD and GADD34) that were up-regulated in patients with progressive disease, suggesting that apoptosis-signaling pathways may be important in chronic sarcoidosis (11).

(Received in original form July 7, 2006; accepted in final form August 26, 2006)

Supported by NIH grants HL71100 (E.S.C.) and HL99-024, HL68019, and HL77732 (D.R.M.).

Correspondence and requests for reprints should be addressed to David R. Moller, M.D., 5501 Hopkins Bayview Circle, Baltimore, MD 21224. E-mail: dmoller@jhmi.edu

Proc Am Thorac Soc Vol 4, pp 101–107, 2007

DOI: 10.1513/pats.200607-140JG

Internet address: www.atsjournals.org

Microarray expression profiling of sarcoidosis BAL cells has been reported by Thonhofer and coworkers using cDNA microarray analysis (12). To test the hypothesis that mycobacterial antigens might act as triggers to the inflammation of sarcoidosis, they studied RNA expression induced by inactivated *Mycobacterium avium* in BAL cells from six untreated patients with active pulmonary sarcoidosis and five control patients with granulomatous disease (two with active tuberculosis and three with active HP). A majority of the 1,500 differentially expressed genes were found to be common to sarcoidosis, *M. tuberculosis* (Mtb) infection, and HP, including cytokines (IFN- α , IL-1 β , and IL-6), major histocompatibility (MHC) class II genes (DR β and DRw53), and growth factors (insulin-like growth factor 2, T-cell factor 1, V-akt, and prothymosin- α). After *M. avium* stimulation, only four genes were exclusively up-regulated in sarcoidosis BAL cells: fatty acid binding protein 4, avian myeloblastosis viral oncogene homolog-like 2, and two unidentified genes. These results suggest there are more signaling pathways in common than different among these specific granulomatous diseases.

Despite the importance of assessing global gene expression profiles of granulomatous tissues, there are no published studies of microarray analyses of sarcoidosis lung tissue. In pilot studies, the authors obtained frozen surgical lung biopsy samples from the HopGenes Advanced Lung Disease Tissue Repository (<http://www.hopkins-genomics.org>) and processed the tissue for total RNA by standard protocol. High-quality RNA deemed suitable for GeneChip analysis was obtained from 6 of 23 patient samples, including one sample from a patient with fibrocystic (stage IV) sarcoidosis, one with nonspecific interstitial pneumonia (NSIP), and two with emphysema. Gene expression was normalized from expression profiles obtained from two normal lung biopsy samples without pathologic histology. First-pass analysis revealed differential expression in over 300 genes ($\sim 2.5\%$ of total on-chip genes) in the fibrocystic sarcoidosis lung explant, including IFN- γ -related genes (IFN- γ receptor, IFN- γ -inducible protein 16, and IL-18), extracellular matrix proteins (laminin, several collagen species, asporin, and versican), and oxidant-related proteins (metallothionein). These preliminary results suggest that pathways known to be relevant to granulomatous inflammation are detected with these methodologies. Most of the tissue samples (17/23) did not yield mRNA of high enough quality for use in the microarray studies, possibly due to inconsistencies in the collection and preservation of earlier tissue samples. It is now widely acknowledged that establishing methods and timing to minimize sample degradation is critical for quality gene expression data (13, 14). Further comparisons of tissues from patients with sarcoidosis during different stages of disease and with other granulomatous and inflammatory lung diseases are likely to provide a more comprehensive understanding of the pathogenesis of sarcoidosis.

The application of expression profiling to develop tools for clinical phenotyping may be of particular importance in a disease such as sarcoidosis, which is characterized by marked heterogeneity in the manifestations of the disease and widely disparate outcomes. Rutherford and colleagues used microarray expression analysis to identify a set of genes preferentially expressed by seven patients with self-limited (all stage I) disease (IL-1 β , IL-8, GRO- β /- γ , CCR2, CCR5, and CCR6) and a second set of genes preferentially expressed by five other patients with persistent (all stage II or III) disease (HLA-DRB1*1501, DQB1*0602, TNF- α , nuclear factor- κ B, cyclic AMP-responsive element modulator, and CD69) (15). This article illustrates the prospect that the development of a discriminatory gene expression set when combined with genetic profiling might help the clinician to distinguish patients with good prognosis from patients likely to develop progressive disease. The construction of

such discriminatory gene sets is dependent on investigator input on the scope of the dataset and the statistical methodologies used in the analysis.

Proteomic approaches have been applied to help in the understanding of the pathogenesis and clinical heterogeneity of sarcoidosis. Newer technologies allow for the mass study of cellular and secreted proteins and their posttranscriptional modifications that complement the results of genomic analyses (16). Proteomic technologies may also be advantageous for sampling lung tissue proteins and proteins that leach into adjacent BAL fluid from healthy tissue or plasma/serum; the ease with which the latter compartments can be sampled provides an opportunity to identify clinically useful biomarkers (17).

Studies have revealed that the protein composition of BAL is complex (18). Initial proteomic studies used two-dimensional (2-D) gel electrophoresis to identify patterns of protein expression followed by selected protein identification by expected gel migration patterns coupled with mass spectrometry (19, 20). Wattiez and coworkers provide data that 2-D gel analysis of BAL samples produces disease-specific fingerprints that may distinguish sarcoidosis from other inflammatory lung diseases (21). Sabouchi-Schütt and coworkers performed 2-D gel electrophoresis analysis of BAL samples coupled with matrix-assisted laser desorption/ionization-time of flight (MALDI-ToF) mass spectrometry peptide analysis to identify 15 differentially expressed proteins in patients with sarcoidosis compared with healthy individuals (22). A follow-up study using sera from the same groups of patients revealed significantly increased expression of 22 proteins in patients with sarcoidosis compared with control individuals. Three of these proteins (immunoglobulin- κ light chain, β_2 microglobulin, β_2 glycoprotein 1) tracked similarly to the levels in the BAL compartment (23). Because this was a hypothesis generating study, further research is necessary to determine whether these proteins represent inflammatory markers or play a pathogenic role in sarcoidosis.

Magi and coworkers performed 2-D gel analysis of BAL samples from six patients with sarcoidosis and six patients with IPF, identifying 32 differentially expressed proteins by MALDI-ToF mass spectrometry (24). Fifteen proteins were up-regulated in sarcoidosis, and 17 proteins were up-regulated in IPF. To demonstrate that this set of proteins may represent a discriminatory profile for these diseases, they extended their analysis with the addition of 10 patients with systemic sclerosis (SSc) and correlated the results with phenotyping of their BAL cells (25). In this study, the cellular profile of SSc BAL cells was found to be intermediate to that of sarcoidosis and IPF, where SSc BAL cells had an intermediate CD4/CD8 lymphocyte ratio compared with sarcoidosis (high % lymphocytes, high CD4/CD8 ratio) and IPF (low % lymphocytes, low CD4/CD8 ratio). The cytokine profile of SSc BAL cells was also intermediate to that of sarcoidosis BAL cells (high IFN- γ expression, low IL-4 expression) and IPF cells (low IFN- γ expression, high IL-4 expression). 2-D gel analysis of SSc BAL cells revealed the expression profile of the 32 previously described proteins to be distinct from that of sarcoidosis and IPF. Determining whether the differential expression of these proteins reflects differences in critical pathogenic pathways specific to these diseases requires further study.

A recent report by Kriegova and coworkers describes correlations of BAL protein expression patterns with clinical phenotyping in sarcoidosis (26). BAL samples from 65 patients with sarcoidosis and 23 healthy individuals were analyzed using protein microarray analysis. They found 40 differentially expressed protein peaks, of which 13 protein peaks were common to patients with Scadding stage I, II, and III chest X-rays (CXR). When individual CXR stages were compared with control subjects, four peaks were differentially expressed in CXR-stage I sarcoidosis,

12 peaks were differentially expressed in CXR-stage II, and 11 peaks were specific to CXR-stage III. Additional comparisons were made between CXR-stage I patients manifesting with or without Löfgren's syndrome (LS). Results from this study demonstrated 25 differentially expressed proteins, 21 of which were up-regulated in LS and four that were up-regulated in non-LS patients. To identify the proteins, three linked procedures were performed: reverse-phase fractionation, sodium dodecyl sulfate-polyacrylamide gel electrophoresis separation, and fingerprint mapping by surface-enhanced laser desorption/ionization-time of flight mass spectra analysis. Three of the 13 BAL proteins associated with sarcoidosis were unambiguously identified: human serum albumin, protocadherin-2 precursor, and α_1 -antitrypsin. Although leaching of these proteins from the plasma due to impairment of lung endothelial barrier function from lung injury seems likely, studies are necessary to determine if protocadherin-2 precursor and α_1 -antitrypsin in sarcoidosis BAL samples are linked with disease phenotype because they play a pathogenic role in different stages of disease. Together, these studies illustrate the potential utility of proteomic profiling for identifying sets of proteins that may provide diagnostic or clinical phenotyping tools to assist in clinical management.

Strategies to improve the sensitivity and specificity of proteomic analyses capitalize on biochemical properties (e.g., pH, charge, size, hydrophobicity) that can be used to limit the proteome and thereby improve the ability to detect and discriminate changes in protein expression. Proteomic profiles can also characterize other molecular properties related to the local inflammatory environment, such as the presence of carbonylated proteins as a marker of oxidative stress (27). Rottoli and coworkers analyzed the presence of carbonylated proteins within the BAL fluid from the above cohort of patients (28). They demonstrated differences in total carbonyl content of BAL protein (sarcoidosis > SSc > IPF) and differences in the species of specific oxidized proteins, suggesting that different pathologic processes are associated with quantitatively different patterns of post-transcriptional protein modifications.

In a previous study, we used a limited proteomic approach to search for candidate pathogenic tissue antigens in sarcoidosis tissues (29). Because granulomas form around a nidus of poorly soluble or insoluble material, we hypothesized that pathobiologically relevant tissue antigens would consist of poorly soluble aggregated proteins but otherwise made no *a priori* assumptions regarding whether potential tissue antigens would be derived from host or microbial proteins. Using extraction protocols to limit the tissue proteome to poorly soluble protein material from sarcoidosis and control tissues, we found a limited number of antigenic protein bands on gel electrophoresis of sarcoidosis tissues. MALDI-ToF mass spectrometry identified one of these antigens to be derived from the Mtb or *M. smegmatis* catalase-peroxidase protein. We found IgG responses to recombinant mycobacterial catalase-peroxidase in more than 50% of patients with sarcoidosis but rarely in purified protein derivative-negative control subjects. This discovery may provide an example that limiting the proteomic (or genomic) set before profiling offers potential advantages by reducing the complexity of the subsequent read-out, if based on pathobiologically relevant, disease-specific observations.

WG

WG and related antineutrophil cytoplasmic autoantibody (ANCA)-associated vasculitides have been the target of genomic and proteomic expression profiling in several studies (30). Csernok and coworkers used PCR to study cytokine gene expression in patients with the ANCA-associated vasculitis of micro-

scopic polyangiitis (MPA) (31). Pathologically and clinically, WG and MPA are differentiated by the presence of granulomatous inflammation in WG. To determine what pathways may be differentially associated with the granulomatous vasculitis, these authors isolated leukocytes from nasal biopsy samples and determined the expression of Th1 and Th2 cytokines by reverse-transcription PCR. They found the preferential expression of IFN- γ in WG and IL-4 in MPA, suggesting that differences in these diseases are partly due to the differential expression of Th1 cytokines, which has been supported by subsequent studies (32).

To gain insight into potential roles for ANCA in WG, Yang and colleagues examined differential gene expression in whole-blood leukocytes that were incubated with proteinase-3-ANCA and myeloperoxidase-ANCA isolated from patients with ANCA-associated vasculitis (33). Analysis by microarray genechip revealed differential expression of several genes induced by ANCA, including differentiation-dependent gene-2, cyclooxygenase-2, and IL-8. Confirming studies demonstrated that differentiation-dependent gene-2 expression in leukocytes correlated strongly with disease activity. Cyclooxygenase-2 expression was also associated with ANCA-associated vasculitis disease activity, whereas IL-8 was associated with disease remission.

In a search for biomarkers that might predict WG disease activity, Stone and colleagues studied 243 serum samples collected from 166 patients with WG using protein microarray analysis followed by surface-enhanced laser desorption/ionization-time of flight fingerprint mass spectra analysis (34). They identified five clusters of proteins that were able to segregate a test group of 72 serum samples according to disease activity with a sensitivity for remission of 95% and a specificity of 91%. Three of the patients who were misclassified as having undergone remission had been treated with intensive immunosuppression before blood draw, and all had elevated clinical scores for disease activity. One of the two serum samples misclassified as active disease was from a patient who was found to have a positive ANCA titer at the time the research sample was obtained but had maintained clinical remission for the subsequent 18 mo. These data offer the possibility that clinically useful biomarkers using serum proteomic profiling may be available soon to assist clinicians in evaluating patients with this devastating disease.

There are no published reports of genomic or proteomic expression profiling of WG lung tissue. Such studies could provide additional insight into the pathogenic mechanisms involved in the diffuse granulomatous inflammation associated with WG and into which pathways may be distinct from other granulomatous lung diseases, such as sarcoidosis, typified by nonvasculitic, compact, epithelioid granulomas (35, 36).

HP

The only study published using microarray analysis of lung tissue from patients with HP involved comparing HP lung tissues with tissues from usual interstitial pneumonia (UIP) and NSIP (37). Selman and colleagues hypothesized that expression profiling could identify specific gene expression "signatures" that may distinguish these histologically overlapping lung disorders. Using gene chip microarray analysis, these investigators analyzed 35 lung tissue samples to construct a discriminatory set of 1,058 genes that included 354 genes up-regulated in UIP relative to HP (largely related with tissue remodeling) and 595 genes up-regulated in HP relative to UIP (largely related with host defense and inflammation). By assigning a score for each biopsy sample based on the relative expression of each of these 1,058 genes, these authors were able to discriminate UIP and HP in all but one case (one HP was reclassified as UIP). When this analysis

was applied to lung tissues from patients with NSIP, these investigators could identify individuals with better (HP-like) and worse (UIP-like) prognosis. One patient reclassified as HP had circulating serum precipitins; two other patients were reclassified as UIP, and a review of their pathology revealed fibrotic NSIP. This study provides an example of how expression profiling could assist in the pathologic phenotyping of these fibrotic lung diseases, although further prospective studies are needed. One limitation in this study is the lack of comparison to healthy lungs or to lungs from other granulomatous diseases to assess gene expression patterns that might provide insight into disease pathogenesis. In this context, given the differences in these diseases, it is not unexpected that a set of genes was found to be differentially expressed.

BERYLLIUM-INDUCED LUNG DISEASE

CBD provides an example of a Th1 cytokine-dominant granulomatous lung disease caused by a known environmental antigen (38, 39). A limited study of two patients using differential display by Gaede and colleagues demonstrated that incubation of PBMCs with beryllium sulfate versus other stimuli (mercury sulfide, lithium carbonate, nickel sulfate, endotoxin, and heat-killed Mtb) resulted in the differential expression of 2.6 to 5.7% of a total of 1,663 sequence tags, suggesting that specific pathways are activated by beryllium salts (40). More recently, Hong-Geller and coworkers used cDNA microarray to study gene expression induced by beryllium in beryllium-naïve PBMCs. They identified up-regulated expression in many inflammation-related genes, including five chemokine genes (Macrophage inhibitory protein [MIP]-1 α , MIP-1 β , growth-related oncogene [GRO]1, GRO3, and RANTES [regulated on activation, T-cell expressed and secreted]) (41). Expression profiling of CBD lung tissues has not been reported but would likely provide important information relevant not only to CBD but also to pulmonary sarcoidosis characterized by similar histopathologic granulomatous inflammation.

MYCOBACTERIAL INFECTION

Mtb and related mycobacterial organisms have been the target of a wealth of genomic and proteomic studies involving the organism and the host response (42). The completion of the genome sequences for Mtb and *M. leprae* and the near completion of other nontuberculous mycobacterial species together with large-scale studies of the proteins expressed by these organisms have provided new insights into the pathogenicity of mycobacterial disease. Comparative analyses of different cultured isolates of Mtb grown under various growth conditions demonstrate how subtle mutations can lead to changes in the Mtb proteome that affect virulence and pathogen survival (43). Expression profiling has also provided insight into mechanisms of drug resistance, such as the identification of gene polymorphisms that are associated with resistance to isoniazid and other Mtb therapies (44, 45). The "selection pressures" of pharmacologic agents and other environmental stimuli have been shown to lead to significant changes in mycobacterial gene expression (46). In one example, Starck and coworkers used 2-D gel electrophoresis to detect differential expression of at least 50 proteins in Mtb (~1% of its resting transcriptome) induced by a change from an anaerobic to an aerobic culture environment (47).

Studies using genomic or proteomic approaches have examined the expression profile of mycobacterial organisms within the intracellular environment of macrophage phagolysosomes. A GeneChip analysis by Schnappinger and colleagues demonstrated that mycobacterial adaptive changes that occur after phagocytosis by host macrophages begin at the transcriptional

level, including the induction of mycobacterial antioxidant and anaerobic fatty acid metabolism pathways, which seemed to be in response to IFN- γ and nitric oxide-associated host responses (48). Monahan and coworkers demonstrated that resident *M. bovis* (Calmette-Guérin bacillus [BCG]) harbored within macrophage-like THP-1 (a human acute monocytic leukemia cell line) cells expressed proteins that were not expressed during typical *in vitro* growth in culture media, including GroEL-1 and GroEL-2, 16 kD α -crystallin, the drug-resistance gene *InhA*, and elongation factor Tu (49). Because the expression profile of mycobacterial organisms may be different within the macrostructures of granulomas that are relevant to active lung disease, Rachman and coworkers performed a comparison study of gene expression of mycobacteria isolated from granulomas in Mtb-infected lung tissue from the pericavity area and more distant lung (50). Using cDNA microarray analysis, they demonstrated unique patterns of mycobacterial gene expression from each of these sites within the Mtb-diseased lung, which were distinct from the gene expression profile of cultured Mtb isolates grown from the same tissues. Together, these studies indicate that the host environment is a critical factor in altering the state of the mycobacterial transcriptome.

Expression profiling of the host response to mycobacterial organisms has been the subject of multiple studies and approaches. For example, microarray analysis of human macrophage gene expression by Wang and colleagues identified the differential expression of Stat-1 and other IFN- γ -related genes that are induced by Mtb infection (51). Using a modification of differential display with cDNA subtraction, Begum and coworkers characterized gene expression in human macrophages in response to *M. bovis* cell wall skeleton and reported patterns of gene expression that were distinct from patterns induced by endotoxin, including the expression of IL-23, zinc-iron transport proteins, and recently described triggering receptor expressed on myeloid cells (TREM) inflammation signal amplifying receptors (52). In a set of microarray experiments, Blumenthal and coworkers identified 50 genes differentially expressed in human macrophages in response to four strains of *M. avium* (including IL-12p40, suppressor of cytokine signalling [SOCS]-1, IL-10, lymphocyte antigen 64, and myosin X), suggesting that the variable human pathogenicity of less virulent mycobacterial strains could also be due to pathogen-host interactions (53).

Because Mtb can be associated with lifetime persistence within infected tissues, it is thought that mycobacteria express proteins that specifically circumvent host immune responses (54). Studies of global macrophage gene expression in response to live Mtb reveal expression profiles distinct from responses to nonpathogenic mycobacteria (*M. smegmatis*) or heat-killed Mtb, including impaired induction of IL-12 and other Th-1-related genes critical for macrophage activation (55). Pai and colleagues determined that prolonged Toll-like receptor signaling induced by multiple Mtb cell wall components may contribute to the impairment of IFN- γ -responsive pathways and other genes related with MHC-II antigen presentation (56, 57). A recent microarray analysis by Mollenkopf and colleagues demonstrated that murine host expression profiles depend on the mode of exposure to a pathogen, with differences found between BCG vaccination, systemic Mtb infection, or aerosolized Mtb lung infection (58). Hisert and coworkers used differential signature-tagged transposon mutagenesis to generate a series of Mtb mutants to identify other Mtb genes (Rv0405, Rv2958c, and Rv0072) that may be involved in countering IFN- γ -dependent immune responses in mice (59).

Animal models of Mtb infection have been used to study strain-specific susceptibility (CBA/J and DBA/2 mice) or resistance (C57BL/6 and BALB/c mice) to Mtb infection. For example, Keller and coworkers performed a microarray analysis to

identify over 100 differentially expressed genes potentially associated with susceptibility or resistance to Mtb in these mouse strains and concluded that susceptibility may be related to over-exuberant cellular inflammation and tissue destruction rather than to a failure to contain infection (60).

Proteomic approaches have been prominently used to assist in the design of vaccines for mycobacterial disease. For example, to identify immunodominant mycobacterial proteins, Covert and colleagues used 2-D gel electrophoresis to isolate subcellular fractions of Mtb cultures, later incubating these fractions with splenocytes isolated from mice previously infected with Mtb, to find 17 novel T-cell antigens (61). A similar approach by Sinha and colleagues using PBMCs isolated from BCG-vaccinated patients to screen Mtb membrane-associated proteins identified several immunodominant proteins that were soluble in Triton X-114 (62). Using an informatics-based approach to screen for immunodominant mycobacterial proteins, Vani and colleagues generated all possible nonoverlapping nonamer peptides from a current list of 52 putative secreted mycobacterial proteins (<http://www.sanger.ac.uk/projects/M.tuberculosis>). They then systematically used the BIMAS algorithm (http://bimas.dcrt.nih.gov/molbio/hla_bind) to predict the likelihood of binding to 33 different HLA-I molecules to assist in the selection of targets for vaccine development (63). A similar bioinformatics approach was used by McMurry and colleagues to identify MHC class II immunodominant epitopes from genomewide scans of Mtb genomes CDC 1551 and H37Rv; they identified 15 epitopes, in particular MT2281–26-J, that induced IFN- γ secretion in PBMCs from a majority of 25 Mtb-immune subjects (64). Information derived from these types of studies may improve the efficiency and effectiveness of a newer generation of vaccines directed toward the prevention of tuberculosis.

LESSONS LEARNED AND FUTURE APPROACHES

Because diseases with granulomatous inflammation share histologic similarities, they likely share common pathogenetic pathways and disease-specific differences. For example, the lessons learned from studies involving genomic and proteomic expression profiling in mycobacterial disease are likely to be relevant to sarcoidosis (65), particularly in light of recent studies that support a mycobacterial etiology in at least a subset of patients with this disease (29). Investigations into how these diseases differ may provide insights into host–pathogen interactions and immunologic responses that may be relevant to HP, WG, and CBD (66). Because of the difficulty in obtaining adequate numbers of tissues from different granulomatous lung diseases and healthy control lung tissues matched for age, gender, tissue location, and smoking status, some studies have used tissues from nongranulomatous lung diseases (e.g., UIP or NSIP) for control groups. These latter comparisons are almost certain to identify differences in gene/protein expression because of the inherent differences in the diseases and may not identify pathways relevant to the pathogenesis of granulomatous lung disease. Collaborative efforts for tissue collection and preservation, such as the Lung Tissue Resource Consortium, may facilitate the organization of larger studies of granulomatous lung disease with appropriate control tissue (<http://grants.nih.gov/grants/guide/rfa-files/RFA-HL-05-005.html>).

One potential hurdle with the use of lung tissue in the study of granulomatous lung diseases is the noise from a bystander effect of surrounding unaffected tissue, which can affect the specificity of genomic expression profiling, as shown for the study of Mtb-infected lung (49). Laser-capture microdissection (LCM) is designed to isolate focal pathologic structures or cells and may be particularly suitable for granulomatous lung disease by

isolating whole granulomas for subsequent genomic and proteomic expression profiling. LCM with real-time PCR has been successfully used in a limited fashion to study the cytokine gene expression in a murine model of Mtb infection (67). The use of LCM may also facilitate comparative analyses of tissue and circulating blood populations to assess how the peripheral blood compartment reflects local pathogenic processes (68).

The technologies involved in the genomic and proteomic approaches discussed in this article are rapidly evolving. Newer technologies, such as proteomic analysis by liquid chromatography coupled with on-line mass spectra analysis, are suited for high throughput analysis of samples and can provide an unbiased sampling of hydrophobic or membrane proteins (69). An integrated approach is needed to construct “response networks” to predict downstream consequences of genomic changes (e.g., proteomic changes and the resultant metabolomic changes) (70). Future improvements in available computational power and the use of multiplex analysis strategies that simultaneously compare multiple source high-throughput data (expression profiles, protein–protein interaction, gene/protein ontology) are likely to improve our ability to predict (novel) gene function and lead to new insights into disease pathogenesis (71, 72).

CONCLUSIONS

Despite the importance of the granulomatous lung diseases as a group, progress has been slow in our understanding of their pathogenesis, predicting their clinical course, and preventing their associated morbidity and mortality. To assist in these goals, expression profiling has been performed using samples from patients with these diseases. The number of studies using gene or protein expression profiling of patients with granulomatous lung disease is limited. Most of these studies examine plasma, serum, or BAL fluid. Expression data using tissues from the different granulomatous diseases are scarce. The reported data produced with these technologies do not allow for any consensus conclusions because independent validation studies are not available. However, these early studies demonstrate the potential of these technologies to provide insights into disease pathogenesis and assist in clinical phenotyping and prediction. Progress has been slowed by the difficulties in obtaining adequate patient samples from the different specific granulomatous lung diseases and by the lack of consensus approaches to the analysis of complex expression profile data. Different strategies may be needed to optimize the generation and analysis of data for different purposes (e.g., the study of disease pathogenesis, clinical phenotyping, or clinical course prediction). Despite these challenges, a collaborative “-omics” approach integrating genomic, proteomic, lipidomic, and metabolomic methods can be envisioned that will fulfill the promise of these technologies to significantly improve our understanding and clinical management of these diseases.

Conflict of Interest Statement: Neither author has a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

References

1. Zhang CG, Chromy BA, McCutchen-Maloney SL. Host-pathogen interactions: a proteomic view. *Expert Rev Proteomics* 2005;2:187–202.
2. Villas-Boas SG, Mas S, Akesson M, Smedsgaard J, Nielsen J. Mass spectrometry in metabolome analysis. *Mass Spectrom Rev* 2005;24:613–646.
3. Hirsch J, Hansen KC, Burlingame AL, Matthay MA. Proteomics: current techniques and potential applications to lung disease. *Am J Physiol Lung Cell Mol Physiol* 2004;287:L1–23.
4. Muller-Quernheim J. Sarcoidosis: immunopathogenetic concepts and their clinical application. *Eur Respir J* 1998;12:716–738.

5. Iida K, Kadota J, Kawakami K, Matsubara Y, Shirai R, Kohno S. Analysis of T cell subsets and beta chemokines in patients with pulmonary sarcoidosis. *Thorax* 1997;52:431-437.
6. Gaede KI, Mamat U, Muller-Quernheim J. Differential gene expression pattern in alveolar macrophages of patients with sarcoidosis and tuberculosis. *J Mol Med* 2004;82:206-210.
7. Gaede KI, Kataria YP, Mamat U, Muller-Quernheim J. Analysis of differentially regulated mRNAs in monocytic cells induced by in vitro stimulation with Kveim-Siltzbach test reagent. *Exp Lung Res* 2004;30:181-192.
8. Wiwien HW, Hiyama K, Maeda H, Ishioka S, Yamakido M. Differential display of messenger RNA expressed in bronchoalveolar lavage cells in pulmonary sarcoidosis patients. *Hiroshima J Med Sci* 1996;45:1-10.
9. Agostini C, Miorin M, Semenzato G. Gene expression profile analysis by DNA microarrays: a new approach to assess functional genomics in diseases. *Sarcoidosis Vasc Diffuse Lung Dis* 2002;19:5-9.
10. Rutherford RM, Kehren J, Staedtler F, Chibout SD, Egan JJ, Tamm M, Gilmartin JJ, Brutsche MH. Functional genomics in sarcoidosis—reduced or increased apoptosis? *Swiss Med Wkly* 2001;131:459-470.
11. Kunitake R, Kuwano K, Miyazaki H, Hagimoto N, Nomoto Y, Hara N. Apoptosis in the course of granulomatous inflammation in pulmonary sarcoidosis. *Eur Respir J* 1999;13:1329-1337.
12. Thonhofer R, Maercker C, Popper HH. Expression of sarcoidosis related genes in lung lavage cells. *Sarcoidosis Vasc Diffuse Lung Dis* 2002;19:59-65.
13. Wang WH, McNatt LG, Shepard AR, Jacobson N, Nishimura DY, Stone EM, Sheffield VC, Clark AF. Optimal procedure for extracting RNA from human ocular tissues and expression profiling of the congenital glaucoma gene FOXC1 using quantitative RTPCR. *Mol Vis* 2001;7:89-94.
14. Baker GL, Shultz MA, Fanucchi MV, Morin DM, Buckpitt AR, Plopper CG. Assessing gene expression in lung subcompartments utilizing in situ RNA preservation. *Toxicol Sci* 2004;77:135-141.
15. Rutherford RM, Staedtler F, Kehren J, Chibout SD, Joos L, Tamm M, Gilmartin JJ, Brutsche MH. Functional genomics and prognosis in sarcoidosis: the critical role of antigen presentation. *Sarcoidosis Vasc Diffuse Lung Dis* 2004;21:10-18.
16. Madi A, Pusztahelyi T, Punyiczki M, Fesus L. The biology of the postgenomic era: the proteomics. *Acta Biol Hung* 2003;54:1-14.
17. Noel-Georis I, Bernard A, Falmagne P, Wattiez R. Proteomics as the tool to search for lung disease markers in bronchoalveolar lavage. *Dis Markers* 2001;17:271-284.
18. Lenz AG, Meyer B, Costabel U, Maier K. Bronchoalveolar lavage fluid proteins in human lung disease: analysis by two-dimensional electrophoresis. *Electrophoresis* 1993;14:242-244.
19. Uebelhoefer M, Bewig B, Oldigs M, Nowak D, Magnussen H, Petermann W, Barth J. Protein profile in bronchoalveolar lavage fluid from patients with sarcoidosis and idiopathic pulmonary fibrosis as revealed by SDS-PAGE electrophoresis and western blot analysis. *Scand J Clin Lab Invest* 1993;53:617-623.
20. Mattow J, Schmidt F, Hohenwarter W, Siejak F, Schaible UE, Kaufmann SH. Protein identification and tracking in two-dimensional electrophoretic gels by minimal protein identifiers. *Proteomics* 2004;4:2927-2941.
21. Wattiez R, Hermans C, Cruyt C, Bernard A, Falmagne P. Human bronchoalveolar lavage fluid protein two-dimensional database: study of interstitial lung diseases. *Electrophoresis* 2000;21:2703-2712.
22. Sabouchi-Schutt F, Astrom J, Hellman U, Eklund A, Grunewald J. Changes in bronchoalveolar lavage fluid proteins in sarcoidosis: a proteomics approach. *Eur Respir J* 2003;21:414-420.
23. Sabouchi-Schutt F, Mikko M, Eklund A, Grunewald J, Astrom J. Serum protein pattern in sarcoidosis analysed by a proteomics approach. *Sarcoidosis Vasc Diffuse Lung Dis* 2004;21:182-190.
24. Magi B, Bini L, Perari MG, Fossi A, Sanchez JC, Hochstrasser D, Paesano S, Raggiaschi R, Santucci A, Pallini V, et al. Bronchoalveolar lavage fluid protein composition in patients with sarcoidosis and idiopathic pulmonary fibrosis: a two-dimensional electrophoretic study. *Electrophoresis* 2002;23:3434-3444.
25. Rottoli P, Magi B, Perari MG, Liberatori S, Nikiforakis N, Bargagli E, Cianti R, Bini L, Pallini V. Cytokine profile and proteome analysis in bronchoalveolar lavage of patients with sarcoidosis, pulmonary fibrosis associated with systemic sclerosis and idiopathic pulmonary fibrosis. *Proteomics* 2005;5:1423-1430.
26. Kriegova E, Melle C, Kolek V, Hutyrova B, Mrazek F, Bleul A, du Bois RM, von Eggeling F, Petrek M. Protein profiles of bronchoalveolar lavage fluid from patients with pulmonary sarcoidosis. *Am J Respir Crit Care Med* 2006;173:1145-1154.
27. Lenz AG, Costabel U, Maier KL. Oxidized BAL fluid proteins in patients with interstitial lung diseases. *Eur Respir J* 1996;9:307-312.
28. Rottoli P, Magi B, Cianti R, Bargagli E, Vagaggini C, Nikiforakis N, Pallini V, Bini L. Carbonylated proteins in bronchoalveolar lavage of patients with sarcoidosis, pulmonary fibrosis associated with systemic sclerosis and idiopathic pulmonary fibrosis. *Proteomics* 2005;5:2612-2618.
29. Song Z, Marzilli L, Greenlee BM, Chen ES, Silver RF, Askin FB, Teirstein AS, Zhang Y, Cotter RJ, Moller DR. Mycobacterial catalase-peroxidase is a tissue antigen and target of the adaptive immune response in systemic sarcoidosis. *J Exp Med* 2005;201:755-767.
30. Kallenberg CG, Brouwer E, Weening JJ, Tervaert JW. Anti-neutrophil cytoplasmic antibodies: current diagnostic and pathophysiological potential. *Kidney Int* 1994;46:1-15.
31. Csernok E, Trabandt A, Muller A, Wang GC, Moosig F, Paulsen J, Schnabel A, Gross WL. Cytokine profiles in Wegener's granulomatosis: predominance of type 1 (Th1) in the granulomatous inflammation. *Arthritis Rheum* 1999;42:742-750.
32. Muller A, Trabandt A, Gloeckner-Hofmann K, Seitzer U, Csernok E, Schonermarck U, Feller AC, Gross WL. Localized Wegener's granulomatosis: predominance of CD26 and IFN-gamma expression. *J Pathol* 2000;192:113-120.
33. Yang JJ, Preston GA, Alcorta DA, Waga I, Munger WE, Hogan SL, Sekura SB, Phillips BD, Thomas RP, Jennette JC, et al. Expression profile of leukocyte genes activated by anti-neutrophil cytoplasmic autoantibodies (ANCA). *Kidney Int* 2002;62:1638-1649.
34. Stone JH, Rajapakse VN, Hoffman GS, Specks U, Merkel PA, Spiera RF, Davis JC, St Clair EW, McCune J, Ross S, et al. A serum proteomic approach to gauging the state of remission in Wegener's granulomatosis. *Arthritis Rheum* 2005;52:902-910.
35. Mark EJ, Matsubara O, Tan-Liu NS, Fienberg R. The pulmonary biopsy in the early diagnosis of Wegener's (pathergic) granulomatosis: a study based on 35 open lung biopsies. *Hum Pathol* 1988;19:1065-1071.
36. DeRemee RA. Sarcoidosis and Wegener's granulomatosis: a comparative analysis. *Sarcoidosis* 1994;11:7-18.
37. Selman M, Pardo A, Barrera L, Estrada A, Watson SR, Wilson K, Aziz N, Kaminski N, Zlotnik A. Gene expression profiles distinguish idiopathic pulmonary fibrosis from hypersensitivity pneumonitis. *Am J Respir Crit Care Med* 2006;173:188-198.
38. Bost T, Newman L, Riches D. Increased TNF-alpha and IL6 mRNA expression by alveolar macrophages in chronic beryllium disease. *Chest* 1993;103:138S.
39. Tinkle S, Kittle L, Schwitters PW, Addison JR, Newman LS. Beryllium stimulates release of T helper 1 cytokines interleukin-2 and interferon gamma from BAL cells in chronic beryllium disease. *Chest* 1996;109:5S-6S.
40. Gaede KI, Mamat U, Schlaak M, Muller-Quernheim J. Analysis of differentially regulated mRNAs in peripheral blood monocytes of berylliosis patients after in vitro stimulation. *J Mol Med* 2000;78:293-299.
41. Hong-Geller E, Pardington PE, Cary RB, Sauer NN, Gupta G. Chemokine regulation in response to beryllium exposure in human peripheral blood mononuclear and dendritic cells. *Toxicology* 2006;218:216-228.
42. Schnappinger D, Schoolnik GK, Ehrt S. Expression profiling of host pathogen interactions: how Mycobacterium tuberculosis and the macrophage adapt to one another. *Microbes Infect* 2006;8:1132-1140.
43. He XY, Zhuang YH, Zhang XG, Li GL. Comparative proteome analysis of culture supernatant proteins of Mycobacterium tuberculosis H37Rv and H37Ra. *Microbes Infect* 2003;5:851-856.
44. Khasnobis S, Escuyer VE, Chatterjee D. Emerging therapeutic targets in tuberculosis: post-genomic era. *Expert Opin Ther Targets* 2002;6:21-40.
45. Hughes MA, Silva JC, Geromanos SJ, Townsend CA. Quantitative proteomic analysis of drug-induced changes in mycobacteria. *J Proteome Res* 2006;5:54-63.
46. Wang R, Prince JT, Marcotte EM. Mass spectrometry of the M. smegmatis proteome: protein expression levels correlate with function, operons, and codon bias. *Genome Res* 2005;15:1118-1126.
47. Starck J, Kallenius G, Marklund BI, Andersson DI, Akerlund T. Comparative proteome analysis of Mycobacterium tuberculosis grown under aerobic and anaerobic conditions. *Microbiology* 2004;150:3821-3829.
48. Schnappinger D, Ehrt S, Voskuil MI, Liu Y, Mangan JA, Monahan IM, Dolganov G, Efron B, Butcher PD, Nathan C, et al. Transcriptional adaptation of Mycobacterium tuberculosis within macrophages: insights into the phagosomal environment. *J Exp Med* 2003;198:693-704.
49. Monahan IM, Betts J, Banerjee DK, Butcher PD. Differential expression of mycobacterial proteins following phagocytosis by macrophages. *Microbiol* 2001;147:459-471.

50. Rachman H, Strong M, Ulrichs T, Grode L, Schuchhardt J, Mollenkopf H, Kosmiadi GA, Eisenberg D, Kaufmann SH. Unique transcriptome signature of *Mycobacterium tuberculosis* in pulmonary tuberculosis. *Infect Immun* 2006;74:1233–1242.
51. 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;39:163–172.
52. Begum NA, Ishii K, Kurita-Taniguchi M, Tanabe M, Kobayashi M, Moriwaki Y, Matsumoto M, Fukumori Y, Azuma I, Toyoshima K, et al. *Mycobacterium bovis* BCG cell wall-specific differentially expressed genes identified by differential display and cDNA subtraction in human macrophages. *Infect Immun* 2004;72:937–948.
53. Blumenthal A, Lauber J, Hoffmann R, Ernst M, Keller C, Buer J, Ehlers S, Reiling N. Common and unique gene expression signatures of human macrophages in response to four strains of *Mycobacterium avium* that differ in their growth and persistence characteristics. *Infect Immun* 2005;73:3330–3341.
54. Huard RC, Chitale S, Leung M, Lazzarini LC, Zhu H, Shashkina E, Laal S, Conde MB, Kritski AL, Belisle JT, et al. The *Mycobacterium tuberculosis* complex-restricted gene *cfp32* encodes an expressed protein that is detectable in tuberculosis patients and is positively correlated with pulmonary interleukin-10. *Infect Immun* 2003;71:6871–6883.
55. McGarvey JA, Wagner D, Bermudez LE. Differential gene expression in mononuclear phagocytes infected with pathogenic and non-pathogenic mycobacteria. *Clin Exp Immunol* 2004;136:490–500.
56. Pai RK, Pennini ME, Tobian AA, Canaday DH, Boom WH, Harding CV. Prolonged toll-like receptor signaling by *Mycobacterium tuberculosis* and its 19-kilodalton lipoprotein inhibits gamma interferon-induced regulation of selected genes in macrophages. *Infect Immun* 2004;72:6603–6614.
57. Fortune SM, Solache A, Jaeger A, Hill PJ, Belisle JT, Bloom BR, Rubin EJ, Ernst JD. *Mycobacterium tuberculosis* inhibits macrophage responses to IFN-gamma through myeloid differentiation factor 88-dependent and -independent mechanisms. *J Immunol* 2004;172:6272–6280.
58. Mollenkopf HJ, Hahnke K, Kaufmann SH. Transcriptional responses in mouse lungs induced by vaccination with *Mycobacterium bovis* BCG and infection with *Mycobacterium tuberculosis*. *Microbes Infect* 2006;8:136–144.
59. Hisert KB, Kirksey MA, Gomez JE, Sousa AO, Cox JS, Jacobs WR Jr, Nathan CF, McKinney JD. Identification of *Mycobacterium tuberculosis* counterimmune (cim) mutants in immunodeficient mice by differential screening. *Infect Immun* 2004;72:5315–5321.
60. 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;84:144–158.
61. Covert BA, Spencer JS, Orme IM, Belisle JT. The application of proteomics in defining the T cell antigens of *Mycobacterium tuberculosis*. *Proteomics* 2001;1:574–586.
62. Sinha S, Kosalai K, Arora S, Namane A, Sharma P, Gaikwad AN, Brodin P, Cole ST. Immunogenic membrane-associated proteins of *Mycobacterium tuberculosis* revealed by proteomics. *Microbiol* 2005;151:2411–2419.
63. Vani J, Shaila MS, Chandra NR, Nayak R. A combined immunoinformatics and structure-based modeling approach for prediction of T cell epitopes of secretory proteins of *Mycobacterium tuberculosis*. *Microbes Infect* 2006;8:738–746.
64. McMurry J, Shai H, Gennaro ML, Carter EJ, Martin W, De Groot AS. Analyzing *Mycobacterium tuberculosis* proteomes for candidate vaccine epitopes. *Tuberculosis (Edinb)* 2005;85:95–105.
65. du Bois R, Goh N, McGrath D, Cullinan P. Is there a role for microorganisms in the pathogenesis of sarcoidosis? *J Intern Med* 2003;253:4–17.
66. Kaufmann SH, Cole ST, Mizrahi V, Rubin E, Nathan C. *Mycobacterium tuberculosis* and the host response. *J Exp Med* 2005;201:1693–1697.
67. Zhu G, Xiao H, Mohan VP, Tanaka K, Tyagi S, Tsen F, Salgame P, Chan J. Gene expression in the tuberculous granuloma: analysis by laser capture microdissection and real-time PCR. *Cell Microbiol* 2003;5:445–453.
68. Jin M, Opalek JM, Marsh CB, Wu HM. Proteome comparison of alveolar macrophages with monocytes reveals distinct protein characteristics. *Am J Respir Cell Mol Biol* 2004;31:322–329.
69. Plymoth A, Yang Z, Lofdahl CG, Ekberg-Jansson A, Dahlback M, Fehniger TE, Marko-Varga G, Hancock WS. Rapid proteome analysis of bronchoalveolar lavage samples of lifelong smokers and never-smokers by micro-scale liquid chromatography and mass spectrometry. *Clin Chem* 2006;52:671–679.
70. Mawuenyega KG, Forst CV, Dobos KM, Belisle JT, Chen J, Bradbury EM, Bradbury AR, Chen X. *Mycobacterium tuberculosis* functional network analysis by global subcellular protein profiling. *Mol Biol Cell* 2005;16:396–404.
71. Joshi T, Chen Y, Becker JM, Alexandrov N, Xu D. Genome-scale gene function prediction using multiple sources of high-throughput data in yeast *Saccharomyces cerevisiae*. *OMICS* 2004;8:322–333.
72. Baty F, Bihl MP, Perriere G, Culhane AC, Brutsche MH. Optimized between-group classification: a new jackknife-based gene selection procedure for genome-wide expression data. *BMC Bioinformatics* 2005;6:239.