

# Microbes Are Associated with Host Innate Immune Response in Idiopathic Pulmonary Fibrosis

Yong Huang<sup>1\*</sup>, Shwu-Fan Ma<sup>1\*</sup>, Milena S. Espindola<sup>2</sup>, Rekha Vij<sup>1</sup>, Justin M. Oldham<sup>3</sup>, Gary B. Huffnagle<sup>4</sup>, John R. Erb-Downward<sup>4</sup>, Kevin R. Flaherty<sup>4</sup>, Beth B. Moore<sup>4</sup>, Eric S. White<sup>4</sup>, Tong Zhou<sup>5</sup>, Jianrong Li<sup>6</sup>, Yves A. Lussier<sup>6</sup>, MeiLan K. Han<sup>4</sup>, Naftali Kaminski<sup>7</sup>, Joe G. N. Garcia<sup>6</sup>, Cory M. Hogaboam<sup>2‡</sup>, Fernando J. Martinez<sup>8‡</sup>, and Imre Noth<sup>1‡</sup>; for the COMET-IPF Investigators

<sup>1</sup>Section of Pulmonary and Critical Care Medicine, Department of Medicine, University of Chicago, Chicago, Illinois; <sup>2</sup>Pulmonary and Critical Care Medicine, Cedars-Sinai Medical Center, Los Angeles, California; <sup>3</sup>Pulmonary and Critical Care Medicine, University of California at Davis, Sacramento, California; <sup>4</sup>Division of Pulmonary and Critical Care Medicine, University of Michigan Health System, Ann Arbor, Michigan; <sup>5</sup>Department of Physiology and Cell Biology, University of Nevada School of Medicine, Reno, Nevada; <sup>6</sup>University of Arizona Health Sciences at the University of Arizona, Tucson, Arizona; <sup>7</sup>Section of Pulmonary, Critical Care and Sleep Medicine, Yale University School of Medicine, New Haven, Connecticut; and <sup>8</sup>Weill Cornell Medical College, Cornell University, New York, New York

## Abstract

**Rationale:** Differences in the lung microbial community influence idiopathic pulmonary fibrosis (IPF) progression. Whether the lung microbiome influences IPF host defense remains unknown.

**Objectives:** To explore the host immune response and microbial interaction in IPF as they relate to progression-free survival (PFS), fibroblast function, and leukocyte phenotypes.

**Methods:** Paired microarray gene expression data derived from peripheral blood mononuclear cells as well as 16S ribosomal RNA sequencing data from bronchoalveolar lavage obtained as part of the COMET-IPF (Correlating Outcomes with Biochemical Markers to Estimate Time-Progression in Idiopathic Pulmonary Fibrosis) study were used to conduct association pathway analyses. The responsiveness of paired lung fibroblasts to Toll-like receptor 9 (TLR9) stimulation by CpG-oligodeoxynucleotide (CpG-ODN) was integrated into microbiome–gene expression association analyses for a subset of individuals. The relationship between associated pathways and circulating leukocyte phenotypes was explored by flow cytometry.

**Measurements and Main Results:** Down-regulation of immune response pathways, including nucleotide-binding oligomerization

domain (NOD)-, Toll-, and RIG1-like receptor pathways, was associated with worse PFS. Ten of the 11 PFS-associated pathways correlated with microbial diversity and individual genus, with species accumulation curve richness as a hub. Higher species accumulation curve richness was significantly associated with inhibition of NODs and TLRs, whereas increased abundance of *Streptococcus* correlated with increased NOD-like receptor signaling. In a network analysis, expression of up-regulated signaling pathways was strongly associated with decreased abundance of operational taxonomic unit 1341 (OTU1341; *Prevotella*) among individuals with fibroblasts responsive to CpG-ODN stimulation. The expression of TLR signaling pathways was also linked to CpG-ODN responsive fibroblasts, OTU1341 (*Prevotella*), and Shannon index of microbial diversity in a network analysis. Lymphocytes expressing C-X-C chemokine receptor 3 CD8 significantly correlated with OTU1348 (*Staphylococcus*).

**Conclusions:** These findings suggest that host–microbiome interactions influence PFS and fibroblast responsiveness.

**Keywords:** host immune response and microbial interaction; peripheral blood mononuclear cell transcriptome; bronchoalveolar lavage microbiome; CpG-oligodeoxynucleotide response; pattern recognition receptors

(Received in original form July 27, 2016; accepted in final form February 2, 2017)

\*These authors contributed equally to the manuscript.

‡Co-senior authors.

Correspondence and requests for reprints should be addressed to Imre Noth, M.D., Section of Pulmonary and Critical Care Medicine, Department of Medicine, The University of Chicago, 5841 South Maryland Avenue, MC6076, Chicago, IL 60637. E-mail: inoth@medicine.bsd.uchicago.edu

This article has an online supplement, which is accessible from this issue's table of contents at [www.atsjournals.org](http://www.atsjournals.org)

Am J Respir Crit Care Med Vol 196, Iss 2, pp 208–219, Jul 15, 2017

Copyright © 2017 by the American Thoracic Society

Originally Published in Press as DOI: 10.1164/rccm.201607-1525OC on February 3, 2017

Internet address: [www.atsjournals.org](http://www.atsjournals.org)

## At a Glance Commentary

### Scientific Knowledge on the

**Subject:** Peripheral blood gene expression profiling has identified immune signaling pathways to be associated with outcomes and disease severity in patients with idiopathic pulmonary fibrosis. Variation in the lower airway microbiome has also been associated with differential outcome risk. Whether the airway microbiome influences host immune response in patients with idiopathic pulmonary fibrosis remains unclear.

### What This Study Adds to the

**Field:** This study demonstrates that host defense, as assessed by immune pathway gene expression, may be modulated by variations in the lower airway microbiome. This study also demonstrates that host–microbiome interaction may influence immune-mediated fibroblast responsiveness and the composition of circulating leukocytes.

Idiopathic pulmonary fibrosis (IPF) is an interstitial lung disease with high mortality and variable disease progression. Whereas some patients progress rapidly to death, others experience a slower decline with periods of stability (1). The causes of IPF and factors leading to disease progression remain incompletely characterized (2).

IPF pathogenesis involves recurrent injury to the alveolar epithelium with aberrant wound healing that results in fibrosis rather than normal repair (3–5). Studies have implicated immunologic aberrations to be associated with IPF progression. These include the development of autoimmunity (6) and dysregulation of immune signaling (7–10). Toll-interacting protein (*TOLLIP*), *TLR3*,

*TLR9*, and *MUC5B* (11–15), all of which contribute to innate immunity and host defense (16–21), have been linked to IPF susceptibility and outcomes. The down-regulation of *CD28*, *ICOS*, *LCK*, and *ITK*, genes involved in T-cell signaling, have also been found to predict transplant-free survival in patients with IPF (9), further supporting an immunologic link. Genes involved in the  $\alpha$ -defensin pathway, critical to host defense, also predict disease severity (22). Importantly, sustained *TLR9* activation in lung myofibroblasts can characterize rapidly progressive IPF (15). Altogether, these data suggest a role for modulation of immune mechanisms in IPF disease activity.

Although the lung has historically been considered sterile, modern culture-independent techniques show diverse populations of bacteria in the lung (23, 24). Microaspiration along with microbial migration, elimination, and relative growth rates determine the composition of the lung microbiome (25, 26). Changes in host microanatomy, cell biology, and innate defenses can alter the dynamics of bacterial turnover, leading to colonization by well-recognized bacterial pathogens. In turn, this dynamic can influence lung inflammation (27).

Aberrant host immunity may lead to colonization and proliferation of pathologic organisms in lower airways, potentially contributing to the recurrent alveolar injury characteristic of IPF pathogenesis. We previously demonstrated that the presence of two operational taxonomic units (OTUs), belonging to *Staphylococcus* and *Streptococcus* spp., in bronchoalveolar lavage (BAL) fluid predicted differential survival in patients with IPF (28). Other investigators also identified *Streptococcus* to be more abundant in patients with IPF than in control subjects and identified increased bacterial burden and decreased diversity to be predictive of poor outcomes in IPF (16). Whether the composition of lower airway microbiome influences IPF-relevant

peripheral blood gene expression pathways remains unknown.

In this investigation, we conducted a comprehensive analysis of host–microbiome interaction by integrating peripheral blood gene expression profiles, lung microbial community, and IPF outcomes. Specifically, we hypothesized that the lung microbiome influences innate and adaptive immune response signaling and that host–microbiome interaction modulates progression-free survival (PFS). We used paired clinical, gene expression, circulating leukocyte, and microbial data derived from patients enrolled in the COMET-IPF (Correlating Outcomes with Biochemical Markers to Estimate Time-Progression in Idiopathic Pulmonary Fibrosis) study (NCT01071707) to address this global hypothesis. In addition, we investigated links between identified lung microbes and CpG-oligodeoxynucleotide (CpG-ODN) responsiveness of lung fibroblasts to innate Toll-like receptor 9 (*TLR9*) activation.

## Methods

### Study Population

COMET-IPF participants were diagnosed using American Thoracic Society/European Respiratory Society criteria (29) and prospectively enrolled at nine U.S. clinical centers. Inclusion and exclusion criteria, along with trial endpoints, were previously described (10, 28, 30). Descriptions of demographics and baseline pulmonary function tests are presented in the online supplement. A composite physiologic index (CPI) (31) was calculated for all participants. PFS, the primary combination endpoint for COMET-IPF, was defined as the time from study enrollment to death, acute exacerbation, lung transplant, or relative change in FVC greater than or equal to 10% or diffusing capacity of the lung for carbon monoxide ( $DL_{CO}$ ) greater than or equal to 15%.

Supported by National Institutes of Health grants R02 HL101740 (F.J.M. and I.N.); HL115618 (B.B.M.); R01 HL123899 (C.M.H.); and U01 HL112696, U01 HL125208, and P01 HL126609 (J.G.N.G.).

Author Contributions: Y.H.: analyzed the data and wrote the manuscript; S.-F.M.: collected data, assisted with data analysis, and prepared and wrote the manuscript; M.S.E.: performed the CpG-oligodeoxynucleotide experiments and assisted with manuscript preparation; R.V. and J.M.O.: contributed to interpretation of results and manuscript preparation; T.Z., J.R.E.-D., G.B.H., J.L., B.B.M., Y.A.L., and C.M.H.: contributed to data analysis; E.S.W., K.R.F., M.K.H., G.B.H., B.B.M., N.K., J.G.N.G., C.M.H., and F.J.M.: contributed to conception and study design, interpretation of the results, and manuscript preparation; I.N.: contributed to conception and study design as well as interpretation of the results, and wrote the manuscript; all authors: reviewed, revised, and approved the manuscript for submission.

### Peripheral Blood Mononuclear Cell Isolation, RNA Extraction, Microarray Hybridization, and Data Processing

RNA was extracted from peripheral blood mononuclear cells (PBMCs) isolated from the blood of participants at study entry for gene expression profiling using a microarray platform. Detailed descriptions of the methods are provided in the online supplement.

### BAL for Microbiome 16S Ribosomal RNA Sequencing and Data Processing

Bronchoscopy was completed at enrollment of patients who were clinically stable and without evidence of active infection, as described previously (28). Descriptions of BAL microbial species determination and data processing are presented in the online supplement.

### Lung Fibroblast Culture

Transbronchial biopsy samples were obtained for fibroblast culture in Dulbecco's modified Eagle medium (Lonza, Walkersville, MD) containing 15% fetal calf serum (Cell Generation, Fort Collins, CO), 100 IU of penicillin, 100 µg/ml streptomycin (Corning, Tewksbury, MA), 292 µg/ml L-glutamine (Corning), and 100 µg/ml Primocin (InvivoGen, San Diego, CA) at 37°C and 10% CO<sub>2</sub>. Fresh medium was added to the fibroblasts every 2–3 days, and the cells were passaged when they were 70 to 90% confluent.

### Data Analysis and Integration

See the online supplement for additional details on data analysis and integration.

**PFS-associated host canonical signaling pathways.** Microarray gene expression data derived from host PBMCs was transformed into a data matrix of pathways samples using the GSVA R/Bioconductor software package (<https://bioconductor.riken.jp/packages/3.1/bioc/html/GSVA.html>), as described in the online supplement. Pathways in the Molecular Signature Database (MSigDB; Broad Institute, Cambridge, MA), a collection of annotated gene sets, were correlated with PFS using Cox proportional hazards regression.

**Correlation of gene coexpression modules with clinical traits and OTUs.** The gene expression profiles of PBMCs were clustered into gene modules on the basis of their coexpression patterns using an unsupervised weighted gene coexpression network analysis (WGCNA) package in R (32). A principal component analysis was

used to calculate an eigengene for each gene module (32, 33). A *P* value less than 0.05 was considered significant for Pearson's correlation among individual gene module eigengenes, clinical traits, and microbial community features.

**Functional pathway enrichment analysis.** Canonical pathways were analyzed using Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems/QIAGEN Bioinformatics, Redwood City, CA). Significant pathways were identified using one-sided Fisher's exact tests. See the online supplement for details.

**Cytoscape network analysis.** Identification and visualization of interaction networks that incorporate clinical traits, canonical pathways, and microbial features were performed using the Cytoscape platform ([www.cytoscape.org](http://www.cytoscape.org)).

**Fibroblast responsiveness to TLR9 stimulation using CpG-ODN.** Transbronchial biopsy specimens were available in a subset of COMET-IPF patients (*n* = 27). Lung fibroblasts derived from these transbronchial biopsies were treated for 24 hours with or without CpG-ODN (15). α-Smooth muscle actin greater than twofold expression (i.e., treated vs. untreated) was deemed responsive. Logistic regression with univariate and multivariate analyses was conducted to determine microbial genera-associated CpG-ODN responsiveness. Support vector machines followed by leave-one-out cross-validation (LOOCV) was used to construct a

microbial model predictive of CpG-ODN responsiveness.

**Correlation of circulating leukocyte phenotypes with microbial measures.** Peripheral blood collected from COMET-IPF patients was available for 32 of the 68 cases and was analyzed for leukocyte phenotypes by flow cytometry as previously described (10).

## Results

### Study Population

Sixty-eight of the 88 subjects enrolled in COMET-IPF possessed both sufficient PBMC RNA and BAL specimens for the paired integration analyses (see Figure E1 in the online supplement). A subset (*n* = 27) of these 68 COMET-IPF subjects had paired CpG-ODN response data for TLR9 stimulation in lung fibroblasts derived from transbronchial biopsies as well as microarray gene expression data derived from host PBMCs (Figure E1). Demographics and clinical traits of these two subsets are summarized in Table 1.

### PFS-associated Canonical Signaling Pathways Are Predominantly Immunologic

The events comprising the composite endpoint in the COMET-IPF cohort (Table E1) are summarized with the Kaplan-Meier estimator (Figure E2). They demonstrated that categorical lung function decline measured by FVC or DL<sub>CO</sub> is the largest

**Table 1.** Baseline Characteristics of Subsets of Patients with Idiopathic Pulmonary Fibrosis in COMET-IPF

| Characteristic                                     | RNA Cases<br>( <i>n</i> = 68)* | Fibroblast Cases<br>( <i>n</i> = 27)† |
|--|--------------------------------|---------------------------------------|
| Sex, male, %                                       | 66.2                           | 77.8                                  |
| Age, yr, mean (SD)                                 | 64.8 (7.2)                     | 64.8 (8.5)                            |
| Height, cm, mean (SD)                              | 170.1 (10.3)                   | 173.7 (8.7)                           |
| Weight, kg, mean (SD)                              | 89.8 (17.6)                    | 93.1 (16.2)                           |
| Progression-free survival at 48 wk, <i>n</i> (%)   | 43 (63.2)                      | 19 (70.4)                             |
| CPI, mean (SD)                                     | 50.5 (11.6)                    | 49.9 (11.7)                           |
| DL <sub>CO</sub> , % predicted baseline, mean (SD) | 43.4 (14.3)                    | 44.7 (14.6)                           |
| FVC, % predicted baseline, mean (SD)               | 70.0 (17.4)                    | 71.5 (17.6)                           |
| FEV <sub>1</sub> , % predicted baseline, mean (SD) | 73.5 (18.9)                    | 76.2 (19.4)                           |
| Gastroesophageal reflux, %                         | 57                             | 63                                    |

**Definition of abbreviations:** COMET-IPF = Correlating Outcomes with Biochemical Markers to Estimate Time-Progression in Idiopathic Pulmonary Fibrosis study; CPI = composite physiologic index; DL<sub>CO</sub> = diffusing capacity of the lung for carbon monoxide.

\*A subset of COMET-IPF patients with expression array data from peripheral blood mononuclear cells, 16S ribosomal RNA sequencing data from bronchoalveolar lavage, and clinical demographics.

†A subset of COMET-IPF patients with CpG-oligodeoxynucleotide-mediated Toll-like receptor 9 response data from fibroblasts and expression array data from peripheral blood mononuclear cells.

constituent of the PFS composite endpoint in this study. At five total events, acute exacerbations were only a minor component. In our pathway analysis of PBMC gene expression microarray data, 11 canonical pathways with odds ratios (ORs) less than 1 were significantly associated with PFS, indicating that inhibition of these pathways was associated with poorer PFS (Table 2). Among these pathways, eight involved the immune inflammatory response and pathogen infection, and three involved pattern recognition receptors (PRRs) responding to pathogen-associated molecular patterns. These included TLRs, nucleotide-binding oligomerization domain (NOD)-like receptors (NODs), and retinoic acid-inducible gene I-like receptor (RIG1) signaling pathways.

### Systems Biology Integration of PFS-associated Pathways and Lung Microbial Community

Ten of 11 PFS-associated pathways (Figure 1, *gray hexagons*) correlated with microbial community features (*green circles*). Network analysis and visualization revealed a microbial feature (species accumulation curve [SAC] richness, represented by *gold circle*) as the hub node in this network, connecting to 7 of 10 pathways with significant negative correlation (*green edges*). The most significant connections were between increased SAC richness and decreased NOD and TLR signaling pathways (*thick*

*green edges*). In addition, OTU1345 (*Streptococcus* genus), found in increased abundance in patients with progressive IPF (34), was negatively (*green edges*) and significantly correlated with the NOD signaling pathway (Figure 1). Increased abundance of OTU1345 (*Streptococcus*), OTU1274 (*Pseudomonas*), OTU1249 (*Leptotrichia*), OTU1350 (*Streptococcus*), and OTU1254 (*Actinomyces*) positively (*red edges*) correlated with PRR pathways (i.e., TLR, NOD2, and RIG1 signaling pathways), which is opposite to SAC richness.

### Gene Coexpression Modules Associated with Clinical Traits, Microbiome Diversity, and Individuals OTUs

A correlation matrix of paired gene expression modules, microbial community features, and clinical traits showed varying degrees of correlation between each of these variables (Figure 2). A list of genes in each module correlated with microbial SAC richness can be found in Table E2. The magenta module was positively correlated with DL<sub>CO</sub> ( $r = 0.32$ ;  $P = 0.007$ ; *red box*) and negatively correlated with CPI ( $r = -0.33$ ;  $P = 0.006$ ; *green box*). This inverse relationship is expected because decreasing DL<sub>CO</sub> leads to increasing CPI.

The magenta module was also positively correlated with several microbial community features, including SAC richness ( $r = 0.52$ ;  $P < 5 \times 10^{-6}$ ), OTU1302 (*Pseudomonadaceae*;  $r = 0.45$ ;  $P < 1 \times 10^{-4}$ ),

OTU1256 (*Prevotella*;  $r = 0.3$ ,  $P = 0.01$ ), and OTU1291 (*Prevotella*;  $r = 0.26$ ;  $P = 0.03$ ). OTU1348 (*Staphylococcus*) and OTU1341 (*Prevotella*) were negatively correlated (i.e.,  $r < 0$ ) with the purple, green, yellow, and pink gene modules. The full taxonomic classification for each OTU is listed in Table E3. A gene significance analysis of the magenta module further confirmed a negative correlation with CPI ( $P = 4.7 \times 10^{-6}$ ) and a positive correlation with microbial richness ( $P = 1.2 \times 10^{-27}$ ) (Figures E3A and E3B, respectively). The canonical pathways within each host gene module are provided in Table E4.

Pathways enriched ( $P < 0.05$ ) within the magenta module included integrin signaling, Rho family GTPase signaling, granulocyte movement, and coagulation system signaling (Figure E4). Integrin  $\alpha_v\beta_3$ , commonly expressed in platelets, is involved in Rho GTPase-mediated activation of sphingosine-1-phosphate endothelial cell chemotaxis (35) and also acts as a receptor for another gene in the set, von Willebrand factor. Sphingosine-1-phosphate activity has been shown to be involved in the onset of pulmonary fibrosis (36, 37). Interestingly, integrin  $\alpha_v\beta_6$ , also linked to pulmonary fibrosis via Rho-mediated signaling (38–40), is not included in the magenta module. Other genes include carbonic anhydrase 2, which is involved in bacterial killing (41), and *PDE5A*, which mediates pulmonary hypertension and has previously been targeted in IPF (42).

### PBMC Canonical Pathways Are Associated with Microbial Community Features

Using significant analysis of microarrays (SAM) with the criterion of false discovery rate (FDR) less than 10%, we identified genes significantly associated with selected microbial community features, including the Shannon index; the inverse Simpson index (invsimpson); SAC richness; and abundance of OTU1348 (*Staphylococcus*), OTU1341 (*Prevotella*), and OTU1302 (*Pseudomonas*) (Table E5). IPA revealed that the integrin-linked kinase signaling pathway and inducible costimulator (iCOS)-iCOS ligand (iCOSL) signaling in T-helper cells were positively and negatively associated with Shannon divergence, respectively. Integrin signaling and epithelial adherence junction signaling pathways were associated with SAC richness index, and natural killer cell signaling was associated with the

**Table 2.** Progression-Free Survival-associated Host Canonical Pathways Identified by Cox Proportional Hazards Regression Analysis

| PFS-associated Signaling Pathways*                           | Odds Ratio | Wald P Value       |
|--|------------|--------------------|
| Toll-like receptor signaling pathway                         | 0.046      | 0.003 <sup>†</sup> |
| RIG1-like receptor signaling pathway                         | 0.022      | 0.005 <sup>†</sup> |
| Regulation of autophagy                                      | 0.027      | 0.005 <sup>†</sup> |
| NOD-like receptor signaling pathway                          | 0.175      | 0.039 <sup>†</sup> |
| Neurotrophin signaling pathway                               | 0.049      | 0.022 <sup>†</sup> |
| Natural killer cell-mediated cytotoxicity                    | 0.035      | 0.009 <sup>†</sup> |
| Epithelial signaling in <i>Helicobacter pylori</i> infection | 0.052      | 0.005 <sup>†</sup> |
| Cytosolic DNA sensing pathway                                | 0.037      | 0.017 <sup>†</sup> |
| Renal cell carcinoma   | 0.073      | 0.015 <sup>†</sup> |
| Adipocytokine signaling pathway                              | 0.038      | 0.017 <sup>†</sup> |
| KEGG chemokine signaling pathway                             | 0.065      | 0.025 <sup>†</sup> |

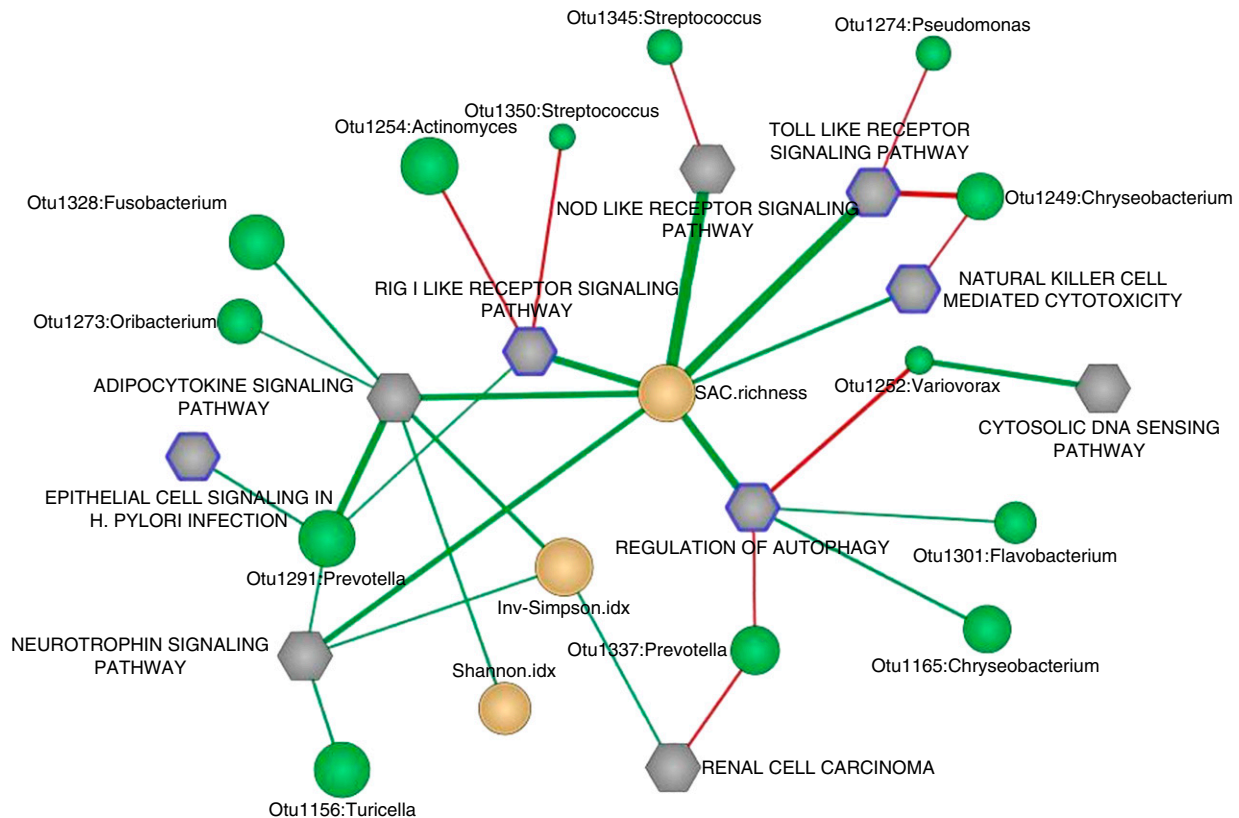
*Definition of abbreviations:* KEGG = Kyoto Encyclopedia of Genes and Genomes; NOD = nucleotide-binding oligomerization domain; PFS = progression-free survival; RIG1 = retinoic acid-inducible gene I-like receptor.

\*Canonical pathways, except the KEGG chemokine signaling pathway, can be found in the interaction network illustrated in Figure 2.

<sup>†</sup> $P < 0.01$ .

<sup>‡</sup> $P < 0.05$ .





**Figure 1.** Interaction network between progression-free survival (PFS)-associated canonical pathways and microbial diversity indices and operational taxonomic units (OTUs). This network illustrates the correlational interaction of PFS-associated host canonical pathways (listed in Table 2), with gray hexagons for pathways with a Wald  $P$  value less than 0.05 or gray hexagons with blue outlines for pathways with a Wald  $P$  value less than 0.01 in PFS analysis, respectively, and microbial community features designated by gold circles for microbial diversity indices and green circles for OTU abundance. The diameter of the green circles is proportional to the correlation coefficient. The red edges represent positive correlation, and the green edges represent negative correlation. The thickness of the edges is determined by  $1 - (P \text{ value})$ . All PFS-associated host canonical pathways except the Kyoto Encyclopedia of Genes and Genomes (KEGG) chemokine signaling pathway in Table 2 correlated with microbial community features. Microbial richness (species accumulation curve index) is the hub node in this network, connecting to 7 of 10 pathways, with significant negative correlation demonstrated by green edges. NOD = nucleotide-binding oligomerization domain; RIG I = retinoic acid-inducible gene I.

invsimpson index. All genes negatively correlated with OTU1348 (*Staphylococcus*) were overrepresented in insulin-like growth factor 1 and vascular endothelial growth factor signaling pathways. All significantly enriched canonical pathways identified by this approach are shown in Table E6. Both WGCNA and SAM approaches supported linkage of diversity index scores with less expression of immune regulatory functions (i.e., iCOS-iCOSL), and more integrin signaling, or cell survival pathways.

#### Microbial Abundance Correlates with TLR Gene Expression in Host PBMCs

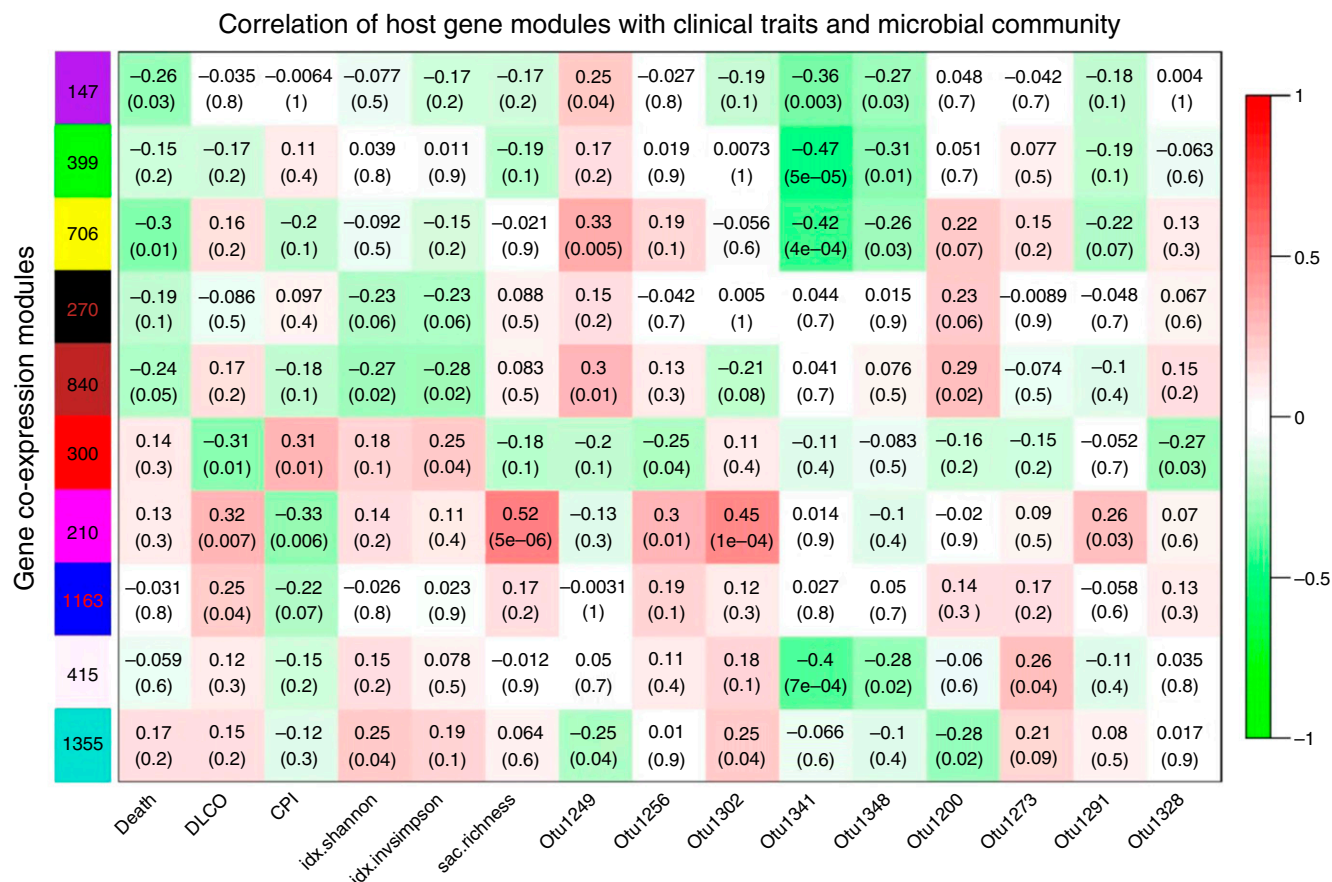
In the list of PFS-associated canonical pathways, TLR signaling was the most prominent feature (Table 2). Thus, we correlated microbial OTUs with expression

of *TLR1–TLR10* and their inhibitory regulator, *TOLLIP* (Figure E5). When correlating TLR mRNA expression levels with the OTUs in BAL fluid, the most significant positive correlation was observed between *TLR9* and OTU1348 (*Staphylococcus*;  $r = 0.38$ ; adjusted  $P$  value [FDR] = 0.02) and OTU1341 (*Prevotella*;  $r = 0.48$ ; FDR = 0.0003). *TOLLIP* expression was correlated with OTU1274 (*Pseudomonas*;  $r = 0.31$ ; FDR = 0.1). OTU1278 (*Bordetella*) was significantly correlated with *TOLLIP* and multiple TLRs, including *TLR2* ( $r = 0.37$ ; FDR = 0.02).

#### Fibroblast Responsiveness to TLR9 Signaling Correlates with Altered Microbial Community

TLR9-mediated CpG-ODN fibroblast responsiveness was associated with a trend

toward differential PFS in this cohort, though the log-rank  $P$  value did not cross the significance threshold, owing in part to the small sample size (Figure E6). Using an iterative support vector machines model followed by LOOCV, we prioritized 10 OTUs that demonstrated 75% sensitivity and 82% specificity in prediction of fibroblast CpG-ODN responsiveness (Table E7). Individual predictors of CpG-ODN responsiveness included OTU1345 (*Staphylococcus*) and OTU1331 (*Veillonella*), which had opposing effect sizes (fold change, 4.07 vs. 0.085, respectively). Notably, four OTUs representing *Prevotella* were identified by this prediction model (Table E7). The abundance of OTU1331 *Veillonella* was significantly correlated with CpG-ODN responsiveness (OR = 1.14; 95% confidence



**Figure 2.** Correlation of gene coexpression modules with clinical traits and microbial community features. Gene coexpression modules were constructed using the WGCNA R software package. The number of genes is labeled within each module. Correlation of the module eigengene with each clinical trait and microbial community features was determined by Pearson's correlation algorithm and is displayed in the corresponding box (coefficient of  $r$  value on top and  $P$  value in parentheses). The color of each box represents the directionality of the correlation (red = positive correlation; green = negative correlation). The bar on the right scales the degree of correlation. Pearson's correlation was used to determine the significance of correlation ( $P < 0.05$ ) between the eigengenes of individual gene modules with clinical traits, including race, sex, age, FVC,  $DL_{CO}$ , CPI, and prognosis, as well as microbial diversity and OTU abundance. CPI = composite physiologic index;  $DL_{CO}$  = diffusing capacity of the lung for carbon monoxide; idx = index; OTU = operational taxonomic unit; sac = species accumulation curves.

interval, 1.02–1.68;  $P = 0.048$ ). The association was maintained after adjustment for sex, age, and CPI (OR, 1.2; 95% confidence interval, 1.01–1.43;  $P = 0.034$ ).

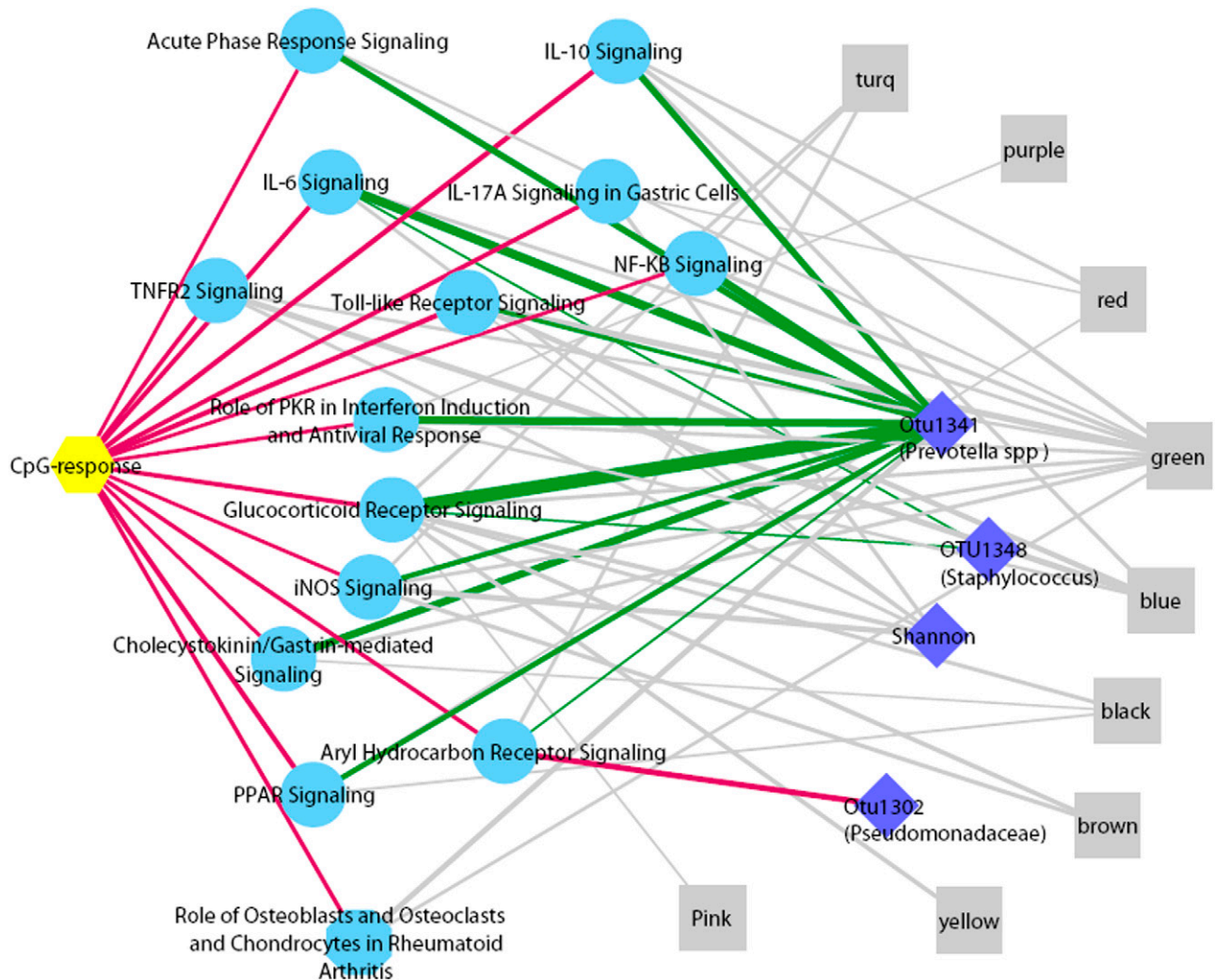
#### Overlap of Canonical Pathways Individually Associated with CpG-ODN Responsiveness in Fibroblasts, Gene Coexpression Modules in Host PBMCs, and Microbial Features Reveals a Network Relationship

We identified 401 differentially expressed genes on the basis of data derived from host PBMCs in cases correlating with CpG-ODN responsiveness in fibroblasts of patients with IPF. An FDR less than or equal to 0.01 was selected to reduce the number of pathways for the Kyoto Encyclopedia of Genes and

Genomes (KEGG) pathway analysis. This analysis revealed that the 93 up-regulated genes associated with CpG-ODN responsiveness were enriched in 22 canonical pathways. These included the immune/inflammatory response and the role of cytokines in mediating communication between immune cells, IL-10, TLR, and IL-6 signaling pathways (Figure E7).

To look at potential interactions, we integrated the 22 CpG-ODN-responsive associated pathways (Figure E7) with pathways enriched within each of the 10 host gene coexpression modules (Table E4), as well as with microbiota-associated pathways (Table E6). We identified 14 overlapping canonical pathways, which were used to construct an interaction

network using Cytoscape version 3.0 software (Figure 3). This network demonstrates that patients with CpG-ODN-responsive fibroblasts (yellow hexagon) are positively correlated (red lines) with numerous immune response pathways (light blue circles). There was a negative association (green lines) between these immune response pathways and the abundance of OTU1341 (*Prevotella*; purple diamond), as well as to, a lesser degree, OTU1348 (*Staphylococcus*; purple diamond). The top four nodes displaying the highest network betweenness centrality were OTU1341 (*Prevotella*), glucocorticoid receptor (GR) signaling, green gene module, and Shannon diversity index (Figure E8A). Three nodes displayed the high closeness centrality, including nuclear



**Figure 3.** Network of canonical signaling pathways that are individually associated with CpG-ODN responsiveness, host gene coexpression modules, and microbial community features. An integrated approach was used to identify common canonical pathways (light blue circles) enriched from the following three paired datasets obtained from patients with IPF in the COMET-IPF study: (1) fibroblast response to CpG-ODN, (2) microbial diversity indices and OTUs, and (3) WGCNA host gene coexpression modules. Up-regulated genes between cases with CpG-ODN response and CpG-ODN nonresponse were subjected to Ingenuity Pathway Analysis with a criterion of FDR less than 0.01 to identify 22 significantly CpG-ODN response-associated canonical pathways. These 22 pathways were then matched to the canonical pathways associated with 10 host gene coexpression modules (gray squares) from PBMC microarray and microbial diversity indices and OTUs (purple diamonds) from bacterial 16S ribosomal RNA sequencing (Tables E2 and E4, respectively). Fourteen pathways (light blue circles) were shared among CpG-ODN response (yellow hexagon), microbial features (purple diamonds), and host gene coexpression modules (gray squares). The red and green edges represent positive and negative associations between nodes, respectively. Gray edges represent both positive and negative associations between nodes. The width of the edge is proportional to the value of  $-\log P$  value (i.e., thicker edge is more significant). COMET-IPF = Correlating Outcomes with Biochemical Markers to Estimate Time-progression in Idiopathic Pulmonary Fibrosis; FDR = false discovery rate; iNOS = inducible nitric oxide synthase; IPF = idiopathic pulmonary fibrosis; NF- $\kappa$ B = nuclear factor- $\kappa$ B; ODN = oligodeoxynucleotide; OTU = operational taxonomic unit; PBMC = peripheral blood mononuclear cell; PKR = protein kinase R; PPAR = peroxisome proliferator-activated receptor; TNFR2 = tumor necrosis factor receptor 2; WGCNA = weighted gene coexpression network analysis R software package.

factor- $\kappa$ B signaling, role of osteoblasts and chondrocytes in rheumatoid arthritis, and acute-phase response signaling (Figure E8B).

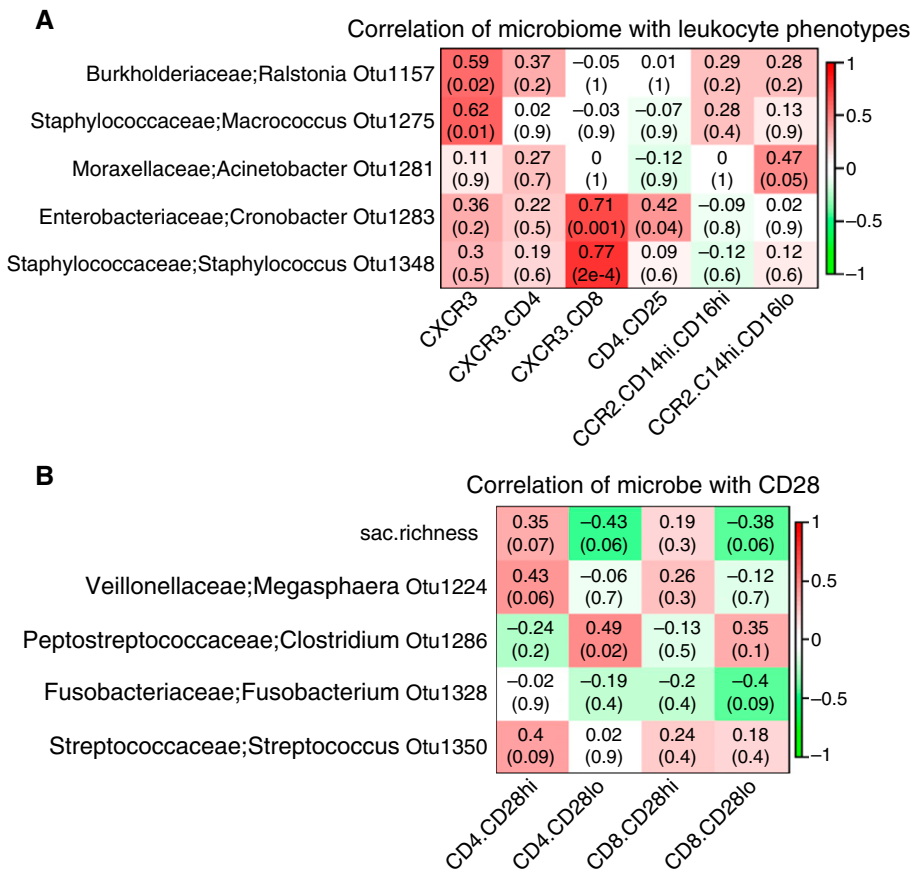
#### Inflammatory Leukocyte Phenotypes Associated with Microbial Measures

Associations between paired lung microbial data and blood leukocyte phenotypes were

assessed in a subset of patients with these data ( $n = 32$ ) (Figure E1). This demonstrated a significant correlation between C-X-C chemokine receptor 3 (CXCR3) CD8-activated T cells and with OTU1348 (*Staphylococcus*;  $r = 0.77$ ; FDR = 0.0002) and OTU1283 (*Cronobacter*;  $r = 0.71$ ; FDR = 0.001) (Figure 4A, red box).

The most significant correlation in CD14-expressing cells was found between CD14<sup>hi</sup>CD16<sup>lo</sup> CC chemokine receptor 2-expressing classical inflammatory macrophages and OTU1281 (*Acinetobacter*;  $r = 0.47$ ; FDR = 0.05) (Figure 4A, red box). Both CD4<sup>+</sup>CD28<sup>lo</sup> and CD8<sup>+</sup>CD28<sup>lo</sup> nonactivated T cells significantly correlated





**Figure 4.** Correlation of circulating leukocyte phenotypes with microbial community. The correlation of microbial diversity indices and OTU abundance with (A) diverse circulating leukocyte phenotypes and (B) CD28 determined by Person's correlation algorithm are shown. To maintain positive or negative directional correlations,  $r$  coefficient instead of  $r^2$  is displayed on top in each box. Criterion for significance is set at coefficient greater than 0.3 (red box) or less than -0.3 (green box) and at FDR less than 0.1 after multiple testing correction shown in parenthesis below each coefficient  $r$  value. CCR = CC chemokine receptor; CXCR = C-X-C chemokine receptor; FDR = false discovery rate; OTU = operational taxonomic unit; sac = species accumulation curve.

with decreased microbial SAC richness ( $r = -0.43$ ; FDR = 0.06; Figure 4B, green box; and  $r = -0.38$ ; FDR = 0.06; Figure 4B, green box) and increased SAC richness correlated with total CD4<sup>+</sup>CD28<sup>hi</sup> nonactivated T cells ( $r = 0.35$ ; FDR = 0.07; Figure 4B, red box).

## Discussion

Microbial pathogenicity, arising from either loss of microbial communal health or loss of an appropriate immune response, has been implicated as a mechanism for variability in IPF disease progression (28). Our comprehensive systems biology approach layered clinical features, including PFS, with multiple biological samples (PBMC gene expression, BAL 16S ribosomal RNA

sequencing, fibroblast CpG-ODN responses, and circulating leukocyte phenotypes) to reveal supportive links to key host immune response pathways. We demonstrate that

1. reduced expression of PFS-associated canonical pathways in PBMCs was associated with a worse outcome (Table 2);
2. down-regulation of immune response-relevant pathways was associated with changes in the abundance of specific microbial OTUs (Figures 1 and 2);
3. abundance of particular OTUs correlated with expression of several TLRs (Figure E5);
4. lung fibroblast CpG-ODN responsiveness to TLR9 activation was

positively correlated with expression of host immune response-relevant signaling pathways (Figure 3);

5. increased expression of these same immune response-relevant pathway genes was negatively correlated with the abundance of OTU1341 (*Prevotella*) and OTU1348 (*Staphylococcus*) (Figure 3); and
6. alterations in the microbial community structure were associated with changes in circulating leukocyte phenotypes (Figure 4).

These data suggest that PRR- and immune response-relevant signaling pathways play a central role in PFS-associated host-microbiome interactions and may provide a target for therapeutic intervention.

In our PFS transcriptomic survival analysis, we identified 11 significant canonical pathways, including major PRR signaling pathways and 4 other inflammatory response pathways (Table 2). TLR-, RIG1-, and NOD-like receptor signaling are integral to the innate immune response, which constitutes the first line of defense against invading microbial pathogens. PRRs detect distinct evolutionarily conserved structures on pathogens, termed "pathogen-associated molecular patterns" (PAMPs) (43). Recognition of PAMPs by PRRs rapidly triggers an array of antimicrobial innate immune responses through the induction of various inflammatory cytokines, chemokines, and type I IFNs to promote host defense.

Nine of the 11, mostly immune, host canonical pathways correlated with both PFS and microbial diversity indices (Figure 1). Most dramatic were increased SAC richness index and a strong negative correlation with NOD-like receptor and TLR signaling pathways. Furthermore, a reduced immune response represented by decreased NOD-like receptor signaling associated with increased abundance of OTU1345 (*Streptococcus*), and both this reduced immune response and increased *Streptococcus* abundance correlated with poorer PFS, suggesting a genomic link. Interestingly, recent work suggests that *Streptococcus* pneumolysin exacerbates lung fibrosis in murine models (44).

Our data also suggest that the microbial community composition is related to impaired epithelial integrity as well as to



severity of lung function abnormality, which is a surrogate of IPF severity. In the WGCNA analysis, bacterial community SAC richness index, OTU1302 (*Pseudomonadaceae*), and preserved  $DL_{CO}$  and CPI were linked to the magenta gene module. By IPA, we found that this module was enriched for integrin and epithelial adherens junction signaling, as well as for signaling by Rho family GTPases and coagulation system, among others (35, 45, 46). All of these are known to be important in epithelial and mesenchymal cell homeostasis (Figure E4) (47–50). Conversely, loss of microbial community diversity is associated with reduced PBMC gene expression in this module. This may be acting as a surrogate for changes of epithelial and mesenchymal cellular activity, which are important in IPF. In addition to OTU1348 (*Staphylococcus*), WGCNA identified OTU1341 (*Prevotella*) as strongly linked to multiple gene expression modules (Figure 2), suggesting a potential symbiosis between *Prevotella* spp. and *Staphylococcus* spp. in reducing PFS.

Examination of genes correlated with quantitative traits by SAM for invsimpson (219 genes) and SAC richness (325 genes) suggested that a more diverse microbial environment leads to an appropriate host immunologic response (Table E5). Conversely, OTU1348 (*Staphylococcus*) significantly correlated with decreased expression of all 464 significant genes identified with overrepresentation in insulin-like growth factor 1 and vascular endothelial growth factor signaling pathways, both of which have been implicated in pulmonary fibrosis (51) and are involved in tissue repair (52).

To our knowledge, our study is the first to link blood gene expression with the lung microbiome. There are studies that have linked the microbiome to lung tissue gene expression (53), as well as PBMC expression and lung tissue expression (54), but none have linked PBMC expression and the microbiome directly. Therefore, any comparison of these prior studies and our present analysis should be undertaken with caution.

Our findings also suggest that the lung fibroblast response to CpG-ODN-mediated TLR9 signaling may be dependent upon the lung microbial community and/or may also reflect the status of immune signaling within the patient as a whole. It is

important to note that, in prior studies by our group using fibroblasts from surgical lung biopsies, only IPF cases demonstrated significant activation by CpG-ODN stimulation, whereas fibroblasts from healthy individuals did not (55). However, these historical surgical biopsies differed in anatomical location from the transbronchial biopsies collected by COMET-IPF investigators, leaving it unknown whether these transbronchial biopsy specimens approximate those that would be observed in surgical lung biopsies. Both alone and within our LOOCV predictor model, increased presence of OTU1331 (*Veillonella*) correlated with increased CpG-ODN fibroblast responsiveness. *Veillonella* is a commensal, anaerobic, gram-negative coccus normally found in the mouth and gut. It commonly triggers the innate immune response via TLR4 (56). Multiple OTUs representing *Prevotella* and *Streptococcus*, which predict reduced PFS in IPF, were also found to be significant negative predictors of fibroblast responsiveness. Positive correlations between TLR9 expression in PBMCs and OTU1348 (*Staphylococcus*) and OTU1341 (*Prevotella*) (Figure E5) were also observed. Although speculative, impaired innate responses in patients with progressive IPF may allow outgrowth of potentially pathogenic species such as *Staphylococcus*, which may in turn promote lung injury and fibrosis. Alternatively, we may be measuring abundance of microbial OTUs (e.g., *Prevotella*) that are not generally considered pathogenic but may cosegregate with other microbial species that are harmful but are not identified in this analysis.

Our network (Figure 3) shows the interrelationship between an altered microbial community, CpG-ODN responsiveness, and PBMC gene expression. Our KEGG analysis of CpG-ODN responsiveness supports activation of the immune/inflammatory response, IL-6 signaling, and TLRs, among others. Furthermore, the magenta module, which is strongly representative of richness, was notably absent in our network analysis. Our network modeling revealed that CpG-ODN responsiveness was connected with OTU1348 (*Staphylococcus*) via GR signaling and with OTU1341 (*Prevotella*) via 10 canonical pathways, including GR, TLR, and IL-6 signaling pathways (Figure 3). This network suggests that

immune responses in PBMCs may be down-regulated in response to accumulation of certain microbial species, depending on fibroblast responsiveness.

Our circulating leukocyte phenotype data provide additional evidence that innate immune responses are aberrant in patients with IPF and may be modulated by alterations in the microbial community. The correlation between a potentially pathogenic genus (*Staphylococcus*; OTU1348) and accumulation of CXCR3<sup>+</sup>CD8<sup>+</sup> T cells involved in Th1 signaling could be evidence that the immune system is activated in response to certain microbes, potentially causing lung damage or, alternatively, that immune cells accumulate to combat the infection. Similarly, we cannot assume that TLR9 signaling is universally beneficial. Although TLR9 activation of professional immune cells may help to limit pathogens, we have also reported that higher expression of TLR9 and consequent CpG-ODN-induced myofibroblast differentiation may accelerate disease progression in IPF.

Our results suggest one possible mechanism for why use of immunosuppressants in the PANTHER-IPF (Prednisone, Azathioprine, and N-Acetylcysteine: A Study That Evaluates Response in Idiopathic Pulmonary Fibrosis) trial proved to be harmful. One patient participated in both PANTHER-IPF and COMET-IPF, though the PANTHER-IPF treatment assignment is unknown. An additional five COMET-IPF patients were treated with low-dose prednisone at the time of bronchoscopy and blood collections, one of whom was also treated with azathioprine. These small numbers precluded the ability to include such therapies in our analysis.

This study has several limitations. Given the study design, our findings are only associative and cannot prove causality. In addition, the expression of some genes identified in this investigation may be influenced by non-PBMC-derived mRNA. This includes *PDE5A* and other vascular system-derived genes. Another limitation stems from the use of a composite endpoint in this analysis. Categorical decline in lung function constituted the largest number of events in this PFS composite endpoint. This may explain why canonical pathways involving innate immunity were highlighted in this investigation, whereas those involved in adaptive immunity with

T-cell signaling were identified as predictive of transplant-free survival previously (9). Limiting the analysis to pulmonary function decline by exclusion of deaths ( $n = 4$ ) and acute exacerbations ( $n = 5$ ) reduced the total number of significant pathways, but it did not appreciably change the results (Table E8).

The size and complexity of this project precluded any reasonable pursuit of a validation cohort. As such, replication of our findings, along with mechanistic confirmation, is needed. Until then, caution is advised when focusing on any individual gene identified by this investigation. We attempted to optimize interpretation of our results by employing several supervised and unsupervised approaches, including GSVA, WGCNA, and network modeling, all of which showed similar findings. We also acknowledge that PBMC transcriptomics and circulating leukocyte phenotypes may not reflect the lung immune response. It is possible that activated inflammatory cells are trafficking to lung, leaving behind peripheral leukocytes that either have not been activated or are unable to be activated. We did not investigate immune cells in the

lung or BAL. Therefore, whether our findings are indicative of lung pathobiology remains to be determined. Furthermore, although the peripheral blood provides an easily accessible compartment for RNA collection, it remains unclear how representative RNA from peripheral blood is of that derived from relevant lung tissue.

A larger-scale replication for blood transcriptomics linked to the bronchoscopic microbiome and clinical outcomes should be pursued in the future and would allow biomarker development for pragmatic clinical application. Additional high-priority investigations include assessing change over time in the microbiome and transcriptomics in the absence and presence of antimicrobial therapy, which would help address the “chicken-and-egg” issue for cause and effect. Last, ascertaining deeper-scale lung microbiome characteristics using whole metagenomic sequencing would allow investigation into microbial functional traits instead of taxa. This has the potential to be significantly more meaningful for pathway analyses and the host immune response.

## Conclusions

Our comprehensive approach sheds light on the molecular mechanism underpinning host–microbiome interaction. In each instance, microbes with increased abundance and decreased community diversity were associated with decreased PBMC transcriptomic expression of immune pathways and poorer PFS. Our data support an evolving concept of pathophysiological fibrosis–like reactions as an essential process in host immune reaction against potential pathogens. These data support the exploration of therapeutic approaches targeting modulation of the lung microbial community of patients with IPF, such as the upcoming National Institutes of Health–sponsored CleanUp-IPF trial (NCT02759120), in which researchers are assessing the efficacy of antibiotic therapy in IPF. More advanced metagenomic analyses are required to elucidate the functional role of individual genera and communities in IPF progression. ■

**Author disclosures** are available with the text of this article at [www.atsjournals.org](http://www.atsjournals.org).

## References

- Nathan SD, Shlobin OA, Weir N, Ahmad S, Kaldjob JM, Battle E, Sheridan MJ, du Bois RM. Long-term course and prognosis of idiopathic pulmonary fibrosis in the new millennium. *Chest* 2011;140:221–229.
- Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, Colby TV, Cordier JF, Flaherty KR, Lasky JA, et al.; ATS/ERS/JRS/ALAT Committee on Idiopathic Pulmonary Fibrosis. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011;183:788–824.
- Kim KK, Kugler MC, Wolters PJ, Robillard L, Galvez MG, Brumwell AN, Sheppard D, Chapman HA. Alveolar epithelial cell mesenchymal transition develops in vivo during pulmonary fibrosis and is regulated by the extracellular matrix. *Proc Natl Acad Sci USA* 2006;103:13180–13185.
- Myers JL, Katzenstein AL. Epithelial necrosis and alveolar collapse in the pathogenesis of usual interstitial pneumonia. *Chest* 1988;94:1309–1311.
- Uhal BD, Joshi I, True AL, Mundle S, Raza A, Pardo A, Selman M. Fibroblasts isolated after fibrotic lung injury induce apoptosis of alveolar epithelial cells in vitro. *Am J Physiol* 1995;269:L819–L828.
- Wilkes DS, Chew T, Flaherty KR, Frye S, Gibson KF, Kaminski N, Klemsz MJ, Lange W, Noth I, Rothhaar K. Oral immunotherapy with type V collagen in idiopathic pulmonary fibrosis. *Eur Respir J* 2015;45:1393–1402.
- Gilani SR, Vuga LJ, Lindell KO, Gibson KF, Xue J, Kaminski N, Valentine VG, Lindsay EK, George MP, Steele C, et al. CD28 down-regulation on circulating CD4 T-cells is associated with poor prognoses of patients with idiopathic pulmonary fibrosis. *PLoS One* 2010;5:e8959.
- O'Dwyer DN, Ashley SL, Moore BB. Influences of innate immunity, autophagy, and fibroblast activation in the pathogenesis of lung fibrosis. *Am J Physiol Lung Cell Mol Physiol* 2016;311:L590–L601.
- Herazo-Maya JD, Noth I, Duncan SR, Kim S, Ma SF, Tseng GC, Feingold E, Juan-Guardela BM, Richards TJ, Lussier Y, et al. Peripheral blood mononuclear cell gene expression profiles predict poor outcome in idiopathic pulmonary fibrosis. *Sci Transl Med* 2013;5:205ra136.
- Moore BB, Fry C, Zhou Y, Murray S, Han MK, Martinez FJ, Flaherty KR; The COMET Investigators. Inflammatory leukocyte phenotypes correlate with disease progression in idiopathic pulmonary fibrosis. *Front Med* 2014;1:56.
- Noth I, Zhang Y, Ma SF, Flores C, Barber M, Huang Y, Broderick SM, Wade MS, Hysi P, Scuirba J, et al. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med* 2013;1:309–317.
- O'Dwyer DN, Armstrong ME, Trujillo G, Cooke G, Keane MP, Fallon PG, Simpson AJ, Millar AB, McGrath EE, Whyte MK, et al. The Toll-like receptor 3 L412F polymorphism and disease progression in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2013;188:1442–1450.
- Fingerlin TE, Murphy E, Zhang W, Peljto AL, Brown KK, Steele MP, Loyd JE, Cosgrove GP, Lynch D, Groshong S, et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet* 2013;45:613–620.
- Peljto AL, Zhang Y, Fingerlin TE, Ma SF, Garcia JG, Richards TJ, Silveira LJ, Lindell KO, Steele MP, Loyd JE, et al. Association between the *MUC5B* promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *JAMA* 2013;309:2232–2239.
- Trujillo G, Meneghin A, Flaherty KR, Sholl LM, Myers JL, Kazerooni EA, Gross BH, Oak SR, Coelho AL, Evanoff H, et al. TLR9 differentiates rapidly from slowly progressing forms of idiopathic pulmonary fibrosis. *Sci Transl Med* 2010;2:57ra82.
- Molyneaux PL, Cox MJ, Willis-Owen SA, Mallia P, Russell KE, Russell AM, Murphy E, Johnston SL, Schwartz DA, Wells AU, et al. The role of bacteria in the pathogenesis and progression of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2014;190:906–913.

17. Shah JA, Vary JC, Chau TT, Bang ND, Yen NT, Farrar JJ, Dunstan SJ, Hawn TR. Human TOLLIP regulates TLR2 and TLR4 signaling and its polymorphisms are associated with susceptibility to tuberculosis. *J Immunol* 2012;189:1737–1746.
18. Saito T, Yamamoto T, Kazawa T, Gejyo H, Naito M. Expression of Toll-like receptor 2 and 4 in lipopolysaccharide-induced lung injury in mouse. *Cell Tissue Res* 2005;321:75–88.
19. Janardhan KS, McIsaac M, Fowlie J, Shrivastav A, Caldwell S, Sharma RK, Singh B. Toll like receptor-4 expression in lipopolysaccharide induced lung inflammation. *Histol Histopathol* 2006;21:687–696.
20. Roy MG, Livraghi-Butrico A, Fletcher AA, McElwee MM, Evans SE, Boerner RM, Alexander SN, Bellinghausen LK, Song AS, Petrova YM, *et al.* Muc5b is required for airway defence. *Nature* 2014;505:412–416.
21. Oldham JM, Ma SF, Martinez FJ, Anstrom KJ, Raghu G, Schwartz DA, Valenzi E, Witt L, Lee C, Vij R, *et al.*; IPFnet Investigators. *TOLLIP*, *MUC5B*, and the response to *N*-acetylcysteine among individuals with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2015;192:1475–1482.
22. Yang IV, Luna LG, Cotter J, Talbert J, Leach SM, Kidd R, Turner J, Kummer N, Kervitsky D, Brown KK, *et al.* The peripheral blood transcriptome identifies the presence and extent of disease in idiopathic pulmonary fibrosis. *PLoS One* 2012;7:e37708.
23. Dickson RP, Erb-Downward JR, Freeman CM, Walker N, Scales BS, Beck JM, Martinez FJ, Curtis JL, Lama VN, Huffnagle GB. Changes in the lung microbiome following lung transplantation include the emergence of two distinct *Pseudomonas* species with distinct clinical associations. *PLoS One* 2014;9:e97214.
24. Suau A, Bonnet R, Sutren M, Godon JJ, Gibson GR, Collins MD, Doré J. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Appl Environ Microbiol* 1999;65:4799–4807.
25. Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. The microbiome and the respiratory tract. *Annu Rev Physiol* 2016;78:481–504.
26. Dickson RP, Huffnagle GB. The lung microbiome: new principles for respiratory bacteriology in health and disease. *PLoS Pathog* 2015;11:e1004923.
27. Huffnagle GB, Dickson RP. The bacterial microbiota in inflammatory lung diseases. *Clin Immunol* 2015;159:177–182.
28. Han MK, Zhou Y, Murray S, Tayob N, Noth I, Lama VN, Moore BB, White ES, Flaherty KR, Huffnagle GB, *et al.*; COMET Investigators. Lung microbiome and disease progression in idiopathic pulmonary fibrosis: an analysis of the COMET study. *Lancet Respir Med* 2014;2:548–556.
29. American Thoracic Society; European Respiratory Society. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. This joint statement of the American Thoracic Society (ATS), and the European Respiratory Society (ERS) was adopted by the ATS board of directors, June 2001 and by the ERS Executive Committee, June 2001. *Am J Respir Crit Care Med* 2002;165:277–304.
30. Naik PK, Bozyk PD, Bentley JK, Popova AP, Birch CM, Wilke CA, Fry CD, White ES, Sisson TH, Tayob N, *et al.*; COMET Investigators. Periostin promotes fibrosis and predicts progression in patients with idiopathic pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol* 2012;303:L1046–L1056.
31. Wells AU, Desai SR, Rubens MB, Goh NS, Cramer D, Nicholson AG, Colby TV, du Bois RM, Hansell DM. Idiopathic pulmonary fibrosis: a composite physiologic index derived from disease extent observed by computed tomography. *Am J Respir Crit Care Med* 2003;167:962–969.
32. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008;9:559.
33. Huang Y, Ma SF, Vij R, Oldham JM, Herazo-Maya J, Broderick SM, Strek ME, White SR, Hogarth DK, Sandbo NK, *et al.* A functional genomic model for predicting prognosis in idiopathic pulmonary fibrosis. *BMC Pulm Med* 2015;15:147.
34. Han MK, Huang YJ, Lipuma JJ, Boushey HA, Boucher RC, Cookson WO, Curtis JL, Erb-Downward J, Lynch SV, Sethi S, *et al.* Significance of the microbiome in obstructive lung disease. *Thorax* 2012;67:456–463.
35. Wang L, Lee JF, Lin CY, Lee MJ. Rho GTPases mediated integrin  $\alpha_v\beta_3$  activation in sphingosine-1-phosphate stimulated chemotaxis of endothelial cells. *Histochem Cell Biol* 2008;129:579–588.
36. Huang LS, Berdyshev E, Mathew B, Fu P, Gorshkova IA, He D, Ma W, Noth I, Ma SF, Pendyala S, *et al.* Targeting sphingosine kinase 1 attenuates bleomycin-induced pulmonary fibrosis. *FASEB J* 2013;27:1749–1760.
37. Huang LS, Berdyshev EV, Tran JT, Xie L, Chen J, Ebenezer DL, Mathew B, Gorshkova I, Zhang W, Reddy SP, *et al.* Sphingosine-1-phosphate lyase is an endogenous suppressor of pulmonary fibrosis: role of S1P signalling and autophagy. *Thorax* 2015;70:1138–1148.
38. Jenkins RG, Su X, Su G, Scotton CJ, Camerer E, Laurent GJ, Davis GE, Chambers RC, Matthay MA, Sheppard D. Ligation of protease-activated receptor 1 enhances  $\alpha_v\beta_6$  integrin-dependent TGF- $\beta$  activation and promotes acute lung injury. *J Clin Invest* 2006;116:1606–1614.
39. Scotton CJ, Krupiczkoj MA, Königshoff M, Mercer PF, Lee YC, Kaminski N, Morser J, Post JM, Maher TM, Nicholson AG, *et al.* Increased local expression of coagulation factor X contributes to the fibrotic response in human and murine lung injury. *J Clin Invest* 2009;119:2550–2563.
40. Xu MY, Porte J, Knox AJ, Weinreb PH, Maher TM, Violette SM, McNulty RJ, Sheppard D, Jenkins G. Lysophosphatidic acid induces  $\alpha_v\beta_6$  integrin-mediated TGF- $\beta$  activation via the LPA2 receptor and the small G protein  $G_{\alpha_q}$ . *Am J Pathol* 2009;174:1264–1279.
41. Tang XX, Fok KL, Chen H, Chan KS, Tsang LL, Rowlands DK, Zhang XH, Dong JD, Ruan YC, Jiang X, *et al.* Lymphocyte CFTR promotes epithelial bicarbonate secretion for bacterial killing. *J Cell Physiol* 2012;227:3887–3894.
42. The Idiopathic Pulmonary Fibrosis Clinical Research Network. A controlled trial of sildenafil in advanced idiopathic pulmonary fibrosis. *N Engl J Med* 2010;363:620–628.
43. Wynn TA. Cellular and molecular mechanisms of fibrosis. *J Pathol* 2008;214:199–210.
44. Knippenberg S, Ueberberg B, Maus R, Bohling J, Ding N, Tort Tarres M, Hoymann HG, Jonigk D, Izykowski N, Paton JC, *et al.* *Streptococcus pneumoniae* triggers progression of pulmonary fibrosis through pneumolysin. *Thorax* 2015;70:636–646.
45. Kudo M, Melton AC, Chen C, Engler MB, Huang KE, Ren X, Wang Y, Bernstein X, Li JT, Atabai K, *et al.* IL-17A produced by  $\alpha\beta$  T cells drives airway hyper-responsiveness in mice and enhances mouse and human airway smooth muscle contraction. *Nat Med* 2012;18:547–554.
46. Campus F, Lova P, Bertoni A, Sinigaglia F, Balduini C, Torti M. Thrombopoietin complements  $G_i$ - but not  $G_q$ -dependent pathways for integrin  $\alpha_{IIb}\beta_3$  activation and platelet aggregation. *J Biol Chem* 2005;280:24386–24395.
47. Horan GS, Wood S, Ona V, Li DJ, Lukashev ME, Weinreb PH, Simon KJ, Hahn K, Allaire NE, Rinaldi NJ, *et al.* Partial inhibition of integrin  $\alpha_v\beta_6$  prevents pulmonary fibrosis without exacerbating inflammation. *Am J Respir Crit Care Med* 2008;177:56–65.
48. Riches DW, Backos DS, Redente EF. ROCK and Rho: promising therapeutic targets to ameliorate pulmonary fibrosis. *Am J Pathol* 2015;185:909–912.
49. Osborn-Heaford HL, Ryan AJ, Murthy S, Racila AM, He C, Sieren JC, Spitz DR, Carter AB. Mitochondrial Rac1 GTPase import and electron transfer from cytochrome c are required for pulmonary fibrosis. *J Biol Chem* 2012;287:3301–3312.
50. Noth I, Anstrom KJ, Calvert SB, de Andrade J, Flaherty KR, Glazer C, Kaner RJ, Olman MA; Idiopathic Pulmonary Fibrosis Clinical Research Network (IPFnet). A placebo-controlled randomized trial of warfarin in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2012;186:88–95.



51. Farkas L, Farkas D, Ask K, Möller A, Gauldie J, Margetts P, Inman M, Kolb M. VEGF ameliorates pulmonary hypertension through inhibition of endothelial apoptosis in experimental lung fibrosis in rats. *J Clin Invest* 2009;119:1298–1311.
52. Lee KS, Park SJ, Kim SR, Min KH, Lee KY, Choe YH, Hong SH, Lee YR, Kim JS, Hong SJ, *et al.* Inhibition of VEGF blocks TGF- $\beta_1$  production through a PI3K/Akt signalling pathway. *Eur Respir J* 2008;31: 523–531.
53. Sze MA, Dimitriu PA, Suzuki M, McDonough JE, Campbell JD, Brothers JF, Erb-Downward JR, Huffnagle GB, Hayashi S, Elliott WM, *et al.* Host response to the lung microbiome in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2015; 192:438–445.
54. Koth LL, Solberg OD, Peng JC, Bhakta NR, Nguyen CP, Woodruff PG. Sarcoidosis blood transcriptome reflects lung inflammation and overlaps with tuberculosis. *Am J Respir Crit Care Med* 2011;184: 1153–1163.
55. Meneghin A, Choi ES, Evanoff HL, Kunkel SL, Martinez FJ, Flaherty KR, Toews GB, Hogaboam CM. TLR9 is expressed in idiopathic interstitial pneumonia and its activation promotes in vitro myofibroblast differentiation. *Histochem Cell Biol* 2008;130:979–992.
56. Matera G, Muto V, Vinci M, Zicca E, Abdollahi-Roodsaz S, van de Veerdonk FL, Kullberg BJ, Liberto MC, van der Meer JW, Focà A, *et al.* Receptor recognition of and immune intracellular pathways for *Veillonella parvula* lipopolysaccharide. *Clin Vaccine Immunol* 2009; 16:1804–1809.