POTENTIAL DIAGNOSTIC AND PROGNOSTIC BIOMARKERS OF IDIOPATHIC PULMONARY FIBROSIS AND THE IMPACT OF E2F8 IN COLLAGEN 1 SYNTHESIS.

by

Louis J. Vuga, MD., M.P.H.

Karl-Franzens University, School of Medicine, Graz, Austria, 1999

MPH, University of Pittsburgh, Graduate of Public Health, 2004

Submitted to the Graduate Faculty of the

School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Louis Justine Vuga

It was defended on

June 7, 2016

and approved by

George D. Leikauf, PhD, Environmental and Occupational Health

Kevin F. Gibson, MD, Medicine

Douglas P. Landsittel, PhD, Medicine

Mauricio Rojas MD, Medicine

Dissertation Director: Galen E. Switzer, PhD, Medicine,

POTENTIAL DIAGNOSTIC AND PROGNOSTIC BIOMARKERS OF IDIOPATHIC PULMONARY FIBROSIS AND THE IMPACT OF E2F8 IN COLLAGEN-1 SYNTHESIS

Louis J. Vuga, PhD, MD., M.P.H.

University of Pittsburgh, 2016

Idiopathic pulmonary Fibrosis (IPF) is a chronic lung disease with median survival about 2-3.5 years (1). Lung transplant is the most effective therapy although pirfenidone (2) and nintetanib (3) delay mortality or disease progression. Biomarkers are used for diagnosis and management of diseases (4) but for IPF, a reliable biomarker is not established (5, 6). In this project, we proposed to investigate lung transcript expressions and plasma proteins as diagnostic or prognostic biomarkers and their implication in Pathogenesis of IPF.

First, we found 44 transcripts qualified after t-test, Benjamini and Hochberg multiple correction, and 1.5-fold change cutoff. We observed 33 transcripts that commonly distinguished IPF from control or COPD. The scoring coefficients of the principal components (COMP) distinguished IPF from control or COPD by using Classification Tree model. The receiver operating characteristic (ROC) curves for scoring coefficients of COMP of IPF vs control was with Area Under Curve (AUC) = 0.90 (95% CI), sensitivity 85% and specificity 91% and IPF or COPD was AUC= 0.88 (95% CI), sensitivity 86 % and specificity 90%.

Second, we measured 24 biomarkers in IPF plasma using Searchlight Protein Array. In cross-sectional study, intercellular adhesion molecule 1 (ICAM), vascular cell adhesion molecule 1 (VCAM1), and S100 calcium binding protein A12 (S100A12), and surfactant protein D (SFTPD) were predictors of IPF mortality ($P < 0.05$). In serial studies, previously identified proteins ICAM1, VCAM1, S100A12, interleukin 8 (IL8) and novel proteins matrix

metallopeptidase 10 (MMP10) and tissue inhibitor of metalloproteinase 1 (TIMP1) were predictors of IPF mortality.

Third, we discovered in lung fibroblasts from IPF a deficiency of E2F transcription factor 8 (E2F8), which is eliminated by an F-box protein 16 (FBXO16). Biologically, E2F8 reduces fibroblast proliferation, and E2F8 also decreases collagen 1A1 mRNA synthesis. Hence, we found a model of lung fibrosis focusing on two novel molecular inputs (E2F8 and FBXO16) linked to fibroblast behavior in subjects with IPF.

In summary, our transcriptomic and plasma proteins analyses revealed the utility of biomarker discovery in advancing precision medicine in IPF. Together these finding provide novel insights that may improve IPF diagnosis, prognostic and potential drug discovery.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to thank the members of my dissertation committee; Dr. George D. Leikauf, Dr. Kevin F. Gibson, Dr. Douglas P. Landsittel, Dr. Mauricio Rojas, and Dr. Galen E. Switzer for their constructive suggestions and encouragements that made this project to be finished in timely fashion. I would like to thank Dr. George Leikauf for his mentorship and leadership that inspired me throughout my post-doctoral fellow period and his determination to accept to be the chair and the advisor of my PhD thesis committee. This project has involved a wealth of Data that were collected from institutions across the United States of America that could not be possible without Dr. Kaminski Naftali and his research Team at Simmons Center for Interstitial Diseases (ILD) at UPMC, University of Pittsburgh. I thank them all for the excellent work and collaborations.

Who has a balanced and healthy life in academic career today? At this moment, I would like to thank my power station that energize me and gave me the spirit to continue joyously when things get odd and wizard. I am overwhelmed with the love that Lina and I received from the south Sudanese community in Pittsburgh and Erie Pennsylvania. It is/was their commitment to make sure that Lina and I succeed in our endeavor. Our Ut-most thanks go to Mr. Benedict Killang and his wife Toyin Ruko Daniel and their children; Hadia, Miracle and Christ for their great comforting Love. Some people are magnetic in our Life; my especial thanks go to Auntie Angelina for introducing me to all her wonderful friends with plenty of experience and wisdoms.

Thank you Mama Keji, all of you are great! A good neighbor is a wish of every one and we were blessed to have uncle Makaryo and Angel as our best neighbor.Lastly, I would like to honor my daughter for being humble and smart child. It needs avillage to raise-up a child. Yes, you did it!

1.0 INTRODUCTION

Idiopathic pulmonary fibrosis (IPF) is a devastating and incurable disease with a ~3 year median survival time (1). IPF affects 185,000 subjects in the United States (7) and pathogenesis of IPF is not well understood. Previously, it was suggested that pulmonary inflammation initiates the process of fibrosis. To date, the paradigm has moved to alveolar epithelial cells malfunction and muddled fibroproliferstion (8). The clinical characteristics of IPF diseases is heterogeneous; stable IPF with slow decline, rapid deterioration, and time of stability intermingled with periods of decline (8). Lung transplantation remains to be the most beneficial therapy(9) although Ofev (nintedanib) and Esbriet (Pirfeniodone) were reported to reduce disease progression among some IPF individuals. No single biomarker is used in clinical setting for diagnosis, disease susceptibility, disease prognosis, disease activity, drug efficacy, and stratification of patients for lung transplant (5, 6, 10). Therefore, novel approaches investigating IPF Biomarkers may deliver potential mechanism of IPF, discovery of novel drugs and will have public health impacts.

The diagnosis of IPF is based on the recommendations of ATS and ERS guidelines for diagnosis (11). IPF has been recognized as a chronic, progressive interstitial lung disease. The Histologic and high resolution computer tomographic scan findings are typically usual interstitial pneumonia (UIP) pattern, comprising of reticular irregularities in subpleural distribution with honeycombing (subpleural cysts). The UIP histology is also described as temporally

heterogeneous area of fibrosis and subepithelial fibroblastic foci. The descripted fibroblastic foci is mainly collection of active proliferative myofibroblasts and fibroblasts(12). IPF is an exclusion diagnosis of other interstitial lung diseases and bronchoscopy procedures are sometimes preformed to collect bronchoalvealar fluid [BALF] in order to determine the etiologies of the diseases although it is not necessary required for diagnosis of IPF. It remains difficult and expensive to make diagnosis of IPF because excluding all other types of interstitial lung diseases is critically important. Moreover, sometimes physicians recommend lung biopsy, which is invasive and very expensive procedure.

In general, biomarkers function either as surrogates for clinical meaningful outcomes that may or may not reveal the mechanism of the disease. For IPF, clinical relevant biomarkers could be for diagnosis, prognostication of mortality, and stratification of patients for surgical procedures. Easily obtainable, measurable, reproducible and reliable biomarkers would serve as perfect markers for serial clinical observation. This is ideal for IPF since the diagnosis is made by using information from radiological imaging and biopsy plus clinical history. The diagnosis of IPF is mainly made by expert physicians in the field of interstitial lung diseases who work in tertiary hospitals and they are not available to many patients due to the geographical location. Therefore peripheral blood biomarkers that can serve as diagnostic tools for IPF would help patients and primary physicians to discover illness earlier. Furthermore, predictive biomarkers could identify the effectiveness of current available drugs and drug in clinical trials to monitor disease activity in a longitudinally study. We focused here on diagnostic and prognostic biomarkers because they are relevant and achievable goals in multicenter research to validate and apply to clinical settings.

Peripheral blood proteins have been studied in the last decades in an attempt to distinguish IPF from control and other diseases. This include matrix metalloptroteinase 7 (MMP7), MMP1 (13), mucin 1, cell surface associated (aka Krebs von den lungen 6) (14) surfactant protein A1 (SFTPA1) (15), alpha defensins (16), and secreted phosphoprotein 1 (SPP1 aka osteopontin) (17). Unfortunately, many of these markers are also increased in the blood of patients with other types of interstitial diseases(18). A single diagnostic biomarker may not be accurate due to several confounding factors like smoking; instead we suggest a panel of biomarkers may perform better in IPF diagnosis in the clinical settings.

Prognostic biomarkers are either genes or proteins that predict outcomes and have had extensive study in IPF. The blood proteins suggested as prognostic biomarkers include: fibulin 1 (19), MUC1 (20), C-X-C motif chemokine ligand 13 (CXCL13) (21), MUC5B (22), lysyl oxidase like 2 (LOXL2) (23) periostin (POSTN) (24), chitinase 3 like 1 (CHI3L1 aka YKL40) (25), MMP1, C-C motif chemokine ligand 18 (CCL18), SFTPA1, and SFTPD (15). Blood cells that have been identified as predictors of mortality in IPF subjects include: CD4(+)CD28(null) T cells (26), CD4(+)CD25(+)FOXP3(+) regulatory T cells (27), semaphorin a+ regulatory T cells (28) and fibrocytes (29). In this current study, we replicated published prognostic biomarkers in peripheral blood such ICAM1, S100A12, SFTPD, and VCAM1 in cross-section studies (30). A novel finding of this project is the discovery of 5 longitudinal biomarkers for IPF by using 250 blood plasma samples in serial collection. Time-dependent Cox proportional hazards model was applied to predict the mortality of IPF subjects. This study revealed reliable and reproducible peripheral blood biomarkers that could be used as prognostic biomarkers for IPF subjects and provided new insights into our understanding of IPF pathogenesis

Collagen 1A1 (COL1A1) and elastin (ELN) are excessively deposited in the IPF lungs compared to healthy subjects (31) (32) and can be regulated by transforming growth factor β1 (TGFB1). E2 transcriptional factor 8 (E2F8) is a member of E2F transcription factors. The E2F protein family is divided into two groups, transcriptional activators (E2F1-3) and transcriptional repressors (E2F4-8). E2Fs engage other factors such as retinoblastoma protein (Rb) or dimerization partner (DP) proteins to control gene transcription (33). E2F1 deficient mice have more biliary fibrosis through early growth response 1 (EGR-1) (a zinc finger protein) (34), but overexpression of E2F1 inhibits COL2A1 formation in cartilage (35). None of the E2 family members have been studied on pulmonary fibrosis. In this study, we examined the level of E2F8 mRNA and Verified E2F8 protein level in IPF lungs, fibroblasts and bleomycin induced mice lung fibrosis were conducted. We hypothesized that E2F8 is targeted for ubiquitination and proteasomal degradation in human lung fibroblasts (hLF), which is mediated by F-box protein 16 (FBXO16). The loss of E2F8, in turn, leads to increased COL1A1 synthesis in IPF lung. Human lung fibroblasts (hLF) transfected with Myc-tagged E2F8 plasmids encoding E2F8 suppressed fibroblasts proliferation or COL1A1 synthesis. Transfection with a short hairpin RNA (shRNA) constructs directed at E2F8 induced fibroblasts proliferation or COL1A1 synthesis. To examine the promoter activities of COL1A1, hLF were transfected with a plasmid containing a *COL1A1* promoter, luciferases reporter construct (*GLucCOL1A1*), with a E2F8 plasmid, or with E2F8 shRNA. In this project we investigated extra cellular matrix accumulation of collagen 1A1 in the IPF lung and the involvement of transcription factor E2F8 plus FBX016 in the IPF disease.

Taken together, this project suggests a basis for future research on diagnostic or prognostic biomarkers of IPF and implication of E2F8 and FBX016 in IPF pathogenesis.

2.0 PAPER 1: POTENTIAL DIAGNOSTIC BIOMARKERS OF IDIOPATHIC PULMONARY FIBROSIS (IPF)

2.1 INTRODUCTION

IPF is an utmost lethal and refractory type of Usual Interstitial Pneumonias (UIP). IPF is an incurable disease with a survival typically less than 4 years (8). The diagnosis of IPF is made by excluding other interstitial lung diseases (ILDs). Histological findings of IPF are characterized by patchy foci of fibroblast proliferation activities, inflammatory activities, areas of intensive fibrosis (honeycombing), intersperse with areas of healthy lung (36). High Resolution Computer Tomography (HRCT) is useful in diagnosing IPF (9) and predicting mortality (37). Several environmental and genetic factors have been suggested to contribute to IPF risk including cigarette smoking (38), drug exposure, and viral infection (39, 40). In United States about 24 million adults have abnormal lung function (41, 42) and majority of the deaths can be ascribed to chronic obstructive pulmonary disease (COPD) with at least 15,000 deaths can be ascribed to IPF (43, 44).

Transcriptomic methods have been extensively used in cancer investigations to classify known or novel disease phenotypes. Scientists documented types of diffuse large B cell lymphoma and other hematological malignancies (45), foretelling the prognosis of solid tumors, metastases and forecasting efficacy of therapies in malignancies (46-48). These methodologies

have yet to be fully used in the field of pulmonary medicine to study IPF, COPD or other pulmonary diseases. Nevertheless, the latest studies investigating pulmonary diseases with similar techniques using large cohorts have contributed to significant understandings of methodology and outcomes of IPF (49-51).

Therefore, we undertook this study using transcriptomics and bioinformatics modeling to distinguish IPF from control or COPD subjects and to examine the functions of transcriptomic signatures that distinguish IPF from control subjects.

2.2 METHODS

2.2.1 Samples collection for Microarrays

Lung Genomics Research Consortium (LGRC) was formed in 2009 by principal investigators from Boston University, University of Colorado, Dana-Faber Cancer Institute, University of Michigan, Mayo Clinic, National Jewish Health, and University of Pittsburgh. It was funded by National Heart, Lung, and Blood Institute (NHLBI) on behalf of National Institute of Health (NIH) through the recovery and Reinvestment Act (ARRA) of 2009. The samples for this study were obtained from Lung Tissues Research Consortium (LTRC), a project funded by NIH. The clinical and genomic data of LTRC and LGRC is available for scientists to understand the genomics basis of chronic obstructive pulmonary disease (COPD), Idiopathic Pulmonary Fibrosis (IPF) and other interstitial lung diseases (ILD).

2.2.2 **Microarray Data**

We analyzed RNA snap frozen tissues and de-identified demographics, pulmonary function testing, HRCT, pathological diagnosis and clinical information. Samples were placed in ice cold Trizol, total RNA was extracted and used as a template for double stranded cDNA synthesis. RNA quantity was determined by spectrometry (260nm) and its integrity by Agilent Bioanalyzer and the calculated RNA integrity number (RIN>8).

Planning and design of microarray experiments – A limitation in design and planning of large scale microarray experiments is the potential for batch effects. To address this concern, we applied block designs to ensure that control RNA (normal lung) was used in every batch. The unique design of the Agilent arrays (4 samples per slide) and the high throughput of the array scanner (48 slides) reduced samples variability.

Array analysis: After quality control and LOESS normalization, differentially expressed genes was identified by using the extensive tools in the ScoreGenes package (52) as well as Significant Analysis of Microarrays (SAM) (53) applying Bioconductor (54, 55). Sam is a statistical technique used for identifying significant transcripts in a set of microarrays experiments (53). Transcripts data set generated from microarray experiments are input variables as well as response variable (control or disease) or multiclass grouping experiments (e.g. lung cancer, kidney cancer or liver cancer). For example, SAM analyzes a statistic di for every transcript i, considering the power of the association between the transcripts and response variables. It is critical to understand that SAM uses repeated permutations of transcripts data set to distinguish a transcript as significant related to response variable. The investigator will set up a cutoff to determine for significance by changing the delta parameter based to the false positive rate / false

negative rate or 1.5 -2 fold change parameter. To correct for multiple testing, we used the calculated q-value method as an extension of the false discovery rate (FDR) that is the proportion of transcripts likely to have been falsely classified by chance as being significant (56). Q-values was computed using the R package for Q calculation(52). Initial studies focused on differences between IPF and control followed by IPF vs COPD transcripts analysis. The microarray data set used for this research is publicly available in the NCBI/GEO repository, accession number GSE47460 and GSE72967, http://www.ncbi.nlm.nih.gov/geo/. The demographic and characteristics were obtained from the Lung Genomics Research Consortium (LGRC, http://www.ltrcpublic.com/ .

2.2.3 **Statistical Analysis**

We note that even if as few as 100 genes among the 29,232 probes with Entrez Gene IDs in the January 2007 annotations of the Agilent 4x44K microarray platform are differentially expressed with fold change of at least 1.5, the proposed study with 100 COPD patients and 60 IPF patients, is powered at 95%, with local FDR 0.00008. Additionally, we performed the following analysis: (1) The gene expressions data were log transformed, (2) two samples t-test was conducted, (3) $p<0.05$ were corrected using Benjamini and Hochberg technique (57). (4) Transcripts were selected based on fulfilling the critical $p < 0.05$ and 1.5-fold change, (5) Principal component analysis was performed to achieve reduction of dimensionality. Principal components analysis (PCA) is commonly used mathematical algorithm that aims to reduce the big data dimensionality while maintaining the most variations in the dataset(58). PCA utilizes identification of directions

that are described as principal components with maximal variation in the data set. By doing so several variables can be represented by few principal components. Based on the cutoff of cumulative proportions setup, the samples can be visualized and the similarities as well as differences between samples groups can be estimated. In this study 44 principle components where generated, however 8 principle of components were enough to offer 80% accountability for the variability for control vs IPF. While 15 principle components out of 69 principle components offered 80% accountability for the variability for COPD vs IPF.

(6) The prediction coefficient of IPF vs Control, and IPF vs COPD were conducted through logistic regression, (6) The classification tree model discriminates between outcomes (e.g. diseases vs control). The model looks for the choices of the predictive biomarkers and create the spilt that exploit the like-hood of the available dataset. Again within each subsequent subset, the algorithm examines the level of each variables to indicate the best split (nodes). This process continue until all subjects are perfectly discriminated, or the sample size becomes very small to be divided. The classification trees model was used to identify potential combination of the prediction coefficient scores of the COMP that could be used to distinguish IPF from control or COPD. The analysis was conducted using the rpart package (https://www.r-project.org/) for recursive partitioning. The classification performance was assessed using the ROCR package (http://rocr.bioinf.mpi-sb.mpg.de/).

Quality of a biomarker is a matter of concern in biomarkers development. It is essential that a biomarker can accurately distinguish between disease and control subjects. In case the measurement of Biomarkers is positive or negative, there will be true positive rate (TPR)/false positive rate (FPR). Sensitivity and specificity are usually used for TPR as well as 1-FPR

We conducted the ROC analysis by extracting the pruned dataset (The classification trees model were pruned using 0.01). The Y –Axis presents sensitivity and X- Axis present 1- specificity. The line through 450 degree is 50/50 line.

In summary, we setup 80% variability of the transcripts to retain sufficient principal component number. The comparison between IPF vs control or IPF vs COPD were first obtained by using logistic regression model (59) and followed by specific tests to predict probability and classify the diseases. These analyses were conducted using StataCorp LP, (College Station, TX) and R-package (https://www.r-project.org/).

2.3 RESULTS

2.3.1 Demographic and clinical information of subjects IPF and COPD

We extracted 78 IPF subjects, 130 control subjects, 150 COPD subjects that were recruited in LGRC study. The characteristics of the participants are summarized in Table 2.1

Table 2.1: Demographic and subjects characteristic of IPF, COPD, and control.

Definition of abbreviations: FVC = Force Vital Capacity precent predicted, FEV1 = Force Expiratory Volume 1 second precent predicted, and DLCO = Diffusing Capacity for Carbon Monoxide percent predicted.

Table 2-1: Demographic and subjects characterisitics of IPF, COPD and control

Group	FVC (%)	FEV1 (%)	DLCO (%)	Smoke home (%)	Smoke work (%)	Ever smoke (%)	Male (%)	Age-Mean	N
IPF	64.6	71.4	49.5	73.8	54.8	63.5	65.9	64.8	77
COPD	79.8	79.8	61.9	75	73.8	95	58	69.3	150
Control	94.5	94.6	82.9	62.5	45.5	44.6	50	62	130

2.3.2 Increased transcripts in IPF compared to either control or COPD

When IPF was compared to control samples, out of 29232 probes, 15723 transcripts were significantly differentially expressed at $P<0.05$. After Benjamini and Hochberg multiple correction procedure and the assumption of 1.5 fold change cutoff, the number of candidates was reduced to 44 transcripts (Table 2.2). The 5 most increased transcripts were IGFL1, ABCA12, LY6D, SPRR1A, and MMP13. To understand the biological important of the transcripts, we performed the functional analysis with Database for Annotation, Visualization and Integrated Discovery (DAVID) that revealed relevant pathways as described in Table 2.3. To investigate the differential expressed transcripts in IPF vs COPD, we used the same criteria for p-values and fold change cutoff 1.5. There were 69 transcripts that were increased in IPF and have the FC of 1.5 compared to COPD (Table 2.4).

Thirty three common transcripts were able to distinguish IPF from either control or COPD (Table 2.5). They included MMP10 that was recently identified by others and us as biomarkers of IPF (13, 60).

Table 2-2: The increased transcripts in IPF when compared to control.

Definition of abbreviations: t-test =Two paired t-test, CTL = Control subjects, BH*0.05 = Benjamini and Hochberg multiple correction, FC= Fold Change. IPF = Idiopathic pulmonary fibrosis

Symbol	Ttest_CTL_IPF	Rank	BH*(5%)	FC	Symbol	ttest CTL_IPF	Rank	BH*(5%)	FC
IGFL2	1.3E-28	5	8.5E-06	2.2	ADAMTS18	6.5E-16	542	9.2E-05	1.6
ABCA12	8.6E-28	11	1.8E-05	2.1	CR2	2.3E-10	2043	3.4E-04	1.5
LY6D	5.3E-23	63	1.0E04	2.1	ADAMTS16	4.5E-16	518	8.8E-05	1.5
SPRR1A	1.1E-27	12	2.0E-05	2.0	FLJ41200	8.6E-11	1849	3.1E-03	1.5
MMP13	6.3E-22	90	1.0E2	1.9	DSC3	3.1E-14	835	1.4E-04	1.5
GREM1	5.0E-24	40	6.8E-05	1.8	CLCA2	5.5E-11	1774	3.0E-03	1.5
GPR87	8.5E-25	30	5.1E-05	1.8	VSIG1	7.8E-19	259	4.4E-04	1.5
UGT1A8	2.2E-19	213	3.6E-04	1.8	FAM159B	9.7E-11	1873	3.2E-03	1.5
KRT6A	2.3E-15	624	1.0E-04	1.7	FCRL5	2.7E-16	485	8.2E-04	1.5
KRT6C	3.3E-13	1071	1.8E-04	1.7	STRA6	9.3E-21	135	2.3E-04	1.5
UGT1A6	4.1E-18	305	5.2E-04	1.7	KRT5	2.8E-16	489	8.3E-04	1.5
BAAT	1.0E-20	138	2.3E-04	1.7	GREM1	2.9E-22	77	1.3E-03	1.5
TMEM229A	2.5E-28	7	1.1E-05	1.7	FRMD5	5.4E-23	64	1.0E-04	1.5
TMEM215	6.3E-21	127	2.1E04	1.6	CYP27C1	5.3E-17	404	6.9E-04	1.5
TTR	5.3E-16	528	9.0E-05	1.6	PADI1	7.8E-09	2891	4.9E-03	1.5
COL17A1	5.4E-27	15	2.5E-05	1.6	SERPINB4	9.3E-08	3659	6.2E-03	1.5
CPNE4	1.3E-21	103	1.7E04	1.6	CXCL13	5.1E-13	1118	1.9E-03	1.5
SERPINB5	1.3E-15	588	1.0E-03	1.6	OGDHL	1.6E-17	363	6.2E-04	1.5
MMP10	4.3E-14	869	1.4E-03	1.6	SERPINB3	5.6E-08	3489	5.9E-03	1.5
GLB1L3	3.9E-16	509	8.7E-04	1.6	DIO2	5.1E-19	243	4.1E-04	1.5
MMP1	9.2E-27	17	2.9E-05	1.6	FDCSP	2.6E-11	1655	2.8E-03	1.5
CYP24A1	3.1E-15	647	1.1E-03	1.6	GJB4	1.4E-11	1560	2.6E-03	1.5

Table 2-3: Functional analysis of increased transcripts in IPF compared with control.
Two of the 9 functional annotation clusters of the transcripts that distinguished IPF from Control.

Annotation Cluster 1	Enrichment Score: 5.01		Count	P_Value	Benjamini
SP_PIR_KEYWORDS	Secreted	RT	14	1.1E-7	1.1E-5
SP_PIR_KEYWORDS	signal	RT	18	1.6E-7	8.5E-6
UP_SEQ_FEATURE	signal peptide	RT	18	1.6E-7	2.8E-5
GOTERM_CC_FAT	extracellular region	RT	15	6.9E-7	2.6E-5
UP_SEQ_FEATURE	disulfide bond	RT	13	8.8E-5	7.5E-3
SP_PIR_KEYWORDS	disulfide bond	RT	13	1.3E-4	4.8E-3
UP_SEQ_FEATURE	glycosylation site:N-linked (GlcNAc...)	RT	14	1.6E-3	8.6E-2
SP_PIR_KEYWORDS	glycoprotein	RT	14	2.6E-3	4.5E-2
Annotation Cluster 2	**Enrichment Score: 1.93**		**Count**	**P_Value**	**Benjamini**
GOTERM_BP_FAT	response to wounding	RT	6	1.9E-3	4.9E-1
SP_PIR_KEYWORDS	cytokine	RT	4	2.4E-3	4.9E-2
GOTERM_MF_FAT	cytokine activity	RT	4	4.4E-3	1.9E-1
GOTERM_CC_FAT	extracellular region part	RT	7	7.9E-3	1.4E-1
KEGG_PATHWAY	Cytokine-cytokine receptor interaction	RT	4	9.3E-3	1.6E-1
GOTERM_BP_FAT	cell-cell signaling	RT	5	1.9E-2	9.0E-1
GOTERM_CC_FAT	extracellular space	RT	5	3.9E-2	3.1E-1
GOTERM_BP_FAT	regulation of secretion	RT	3	5.0E-2	9.7E-1
GOTERM_BP_FAT	regulation of cellular localization	RT	3	7.3E-2	9.8E-1

Table 2-4: Increased transcripts in IPF when compared with COPD.

The list of 69 transcripts that distinguished IPF from COPD after 1.5 FC cutoff

Symbol	Symbol	Symbol	Symbol
ABCA12	KRT6A	ASB5	YSK4
IGFL2	SPOCK3	C6orf142	PMCH
MMP13	LGSN	ATP12A	C3orf16
LY6D	DSG3	VTCN1	TMEM229A
UGT1A8	CLCA2	CYP27C1	DSC3
GPR87	VSIG1	DNAH12	ADAMTS16
SPRR1A	RIMS2	CCDC67	LRRIQ1
UGT1A6	GLB1L3	STRA6	LEKR1
TTR	BAAT	CCDC39	KLK13
HS6ST3	GJB4	MMP10	PROM1
CPNE4	PCSK1	UCN2	ADAMTS18
KRT6C	DIO2	FRMD5	TFAP2A
TMEM215	GJB5	OGDHL	LOC100131046
FAM159B	DACH2	NELL1	C7orf57
CYP27C1	LOC554202	TMEM212	ANXA13
HHLA2	C1QTNF8	DKFZp434J0226	
COL17A1	CYP24A1	ECT2L	
SERPINB5	CLDN8	KRT5	

Table 2-5: The common transcripts that distinguish IPF from either from control or COPD.

The 33 transcripts that commonly distinguished IPF either from control or COPD subjects after application the reduction methods being used

Symbol	Symbol	Symbol	Symbol
ABCA12	CPNE4	FRMD5	VSIG1
IGFL2	KRT6C	OGDHL	GLB1L3
MMP13	TMEM215	KRT5	BAAT
LY6D	FAM159B	TMEM229A	GJB4
UGT1A8	CYP27C1	DSC3	PCSK1
GPR87	COL17A1	ADAMTS16	DIO2
SPRR1A	SERPINB5	ADAMTS18	
UGT1A6	STRA6	KRT6A	
TTR	MMP10	CLCA2	

2.3.3 The scoring coefficients of principal components distinguish IPF either from control or COPD subject.

To examine whether the scoring coefficients of principal components computed from transcripts expression accurately classify IPF subjects from control or COPD subjects, we used recursive partitioning to the transcript dataset and discovered that scoring coefficients of the principal components distinguish IPF from COPD or control. The classification model analysis demonstrated that COMP1 <0.07 concentration could discriminate control from IPF with a node (93/14) (CONTROL/IPF) with terminal node of (91/9) while COMP1< 0.07 IPF from control (37/64) with terminal node (4/47) (CONTROL/IPF) (Figure 2.1A). The ROC curves for application of scoring coefficients of 8 principal components to categorize the samples as IPF or control was significant with AUC= 90% (95% CI) (Figure 2.1B), sensitivity 85% and specificity 91% .

Finally, after comparing IPF with control, we were curious to examine whether there exist difference between IPF subjects and COPD subjects in their transcript expression pattern. After following the same procedures above, we found that 69 transcript were increased in IPF compared with COPD. Furthermore, the scoring coefficients of the principal components differentiated COPD from IPF were 15 COMPs. The CART analysis revealed that COMP4 >0.02 concentration could discriminate COPD from IPF with terminal node (137/20) with terminal node (135/11) (COPD/IPF) while on right side of the CART, the terminal node (13/58) (COPD/IPF) (Figure 2.2A). The ROC curves for the scoring coefficients of 15 principal components to categorize the samples as IPF or COPD was significant with AUC= 88% (95% CI) (Figure 2.2B) sensitivity 86 % and specificity 90%.

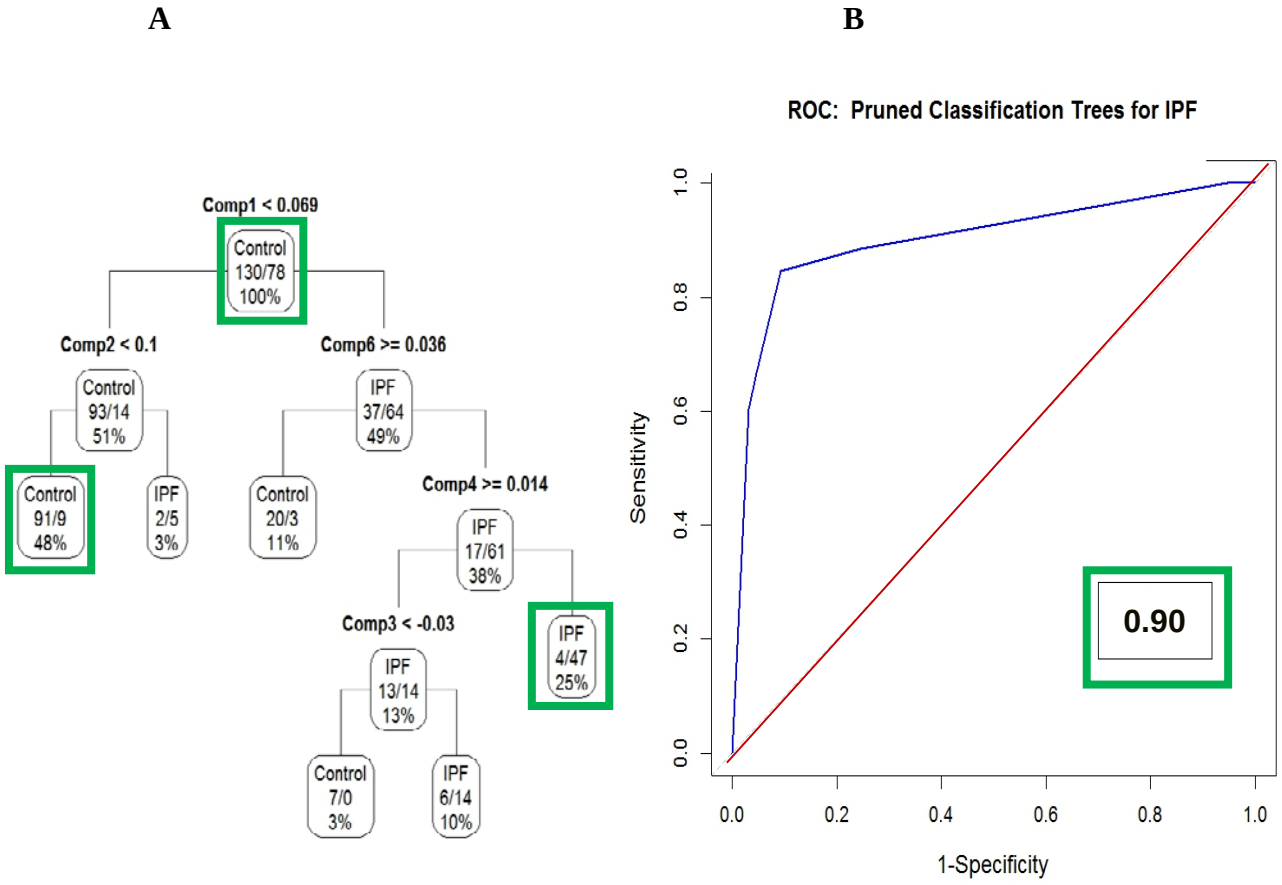


Figure 2-1: Classification tree model and ROC Curve

A: Classification tree model generated by using scoring coefficients of 8 principal components for 44 IPF transcript expression and control transcript expression. Control samples is at the left and IPF in the right (control/IPF) in the box. The numbers counts are control / IPF. The original tree has been pruned with 0.01 value of the scoring coefficients to improve the classification tree. The terminal nodes showed how accurate the Control or IPF were classified. B: The ROC curves for application of scoring coefficients of 8 principal components to categorize the samples as IPF or control. On the x–axis is plotted false positive rate (1- specificity) and y–axis is

sensitivity plotted. The area under curve is corresponding to the numerator of the Mann-Whitney U-statistic matching the scoring coefficients disseminations between IPF and control. The red Line at 45 degree represents 50% classification of IPF vs control participants.

A **B**

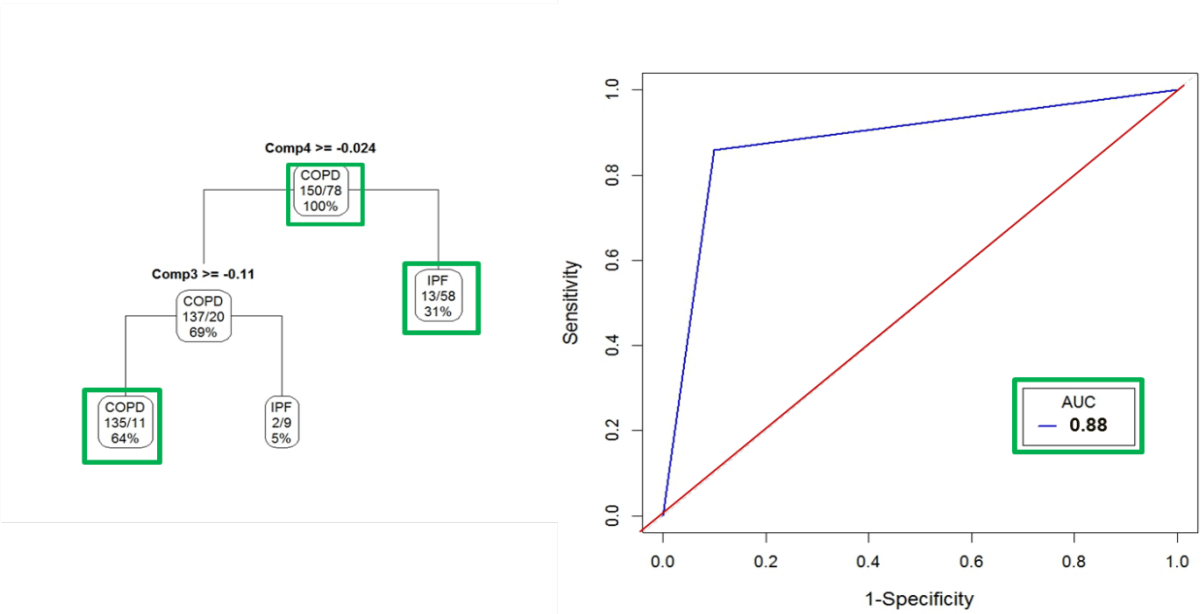


Figure 2-2: Classification tree model and ROC Curve

A: Classification tree model generated by using to scoring coefficients of 15 principal components for 69 IPF transcripts expression and COPD transcript expression. COPD samples is at the left and IPF in the right (COPD/IPF) in the box. The numbers counts are COPD / IPF. The original tree has been pruned with 0.01 values of the scoring coefficients to improve the classification tree. The terminal nodes indicate how accurate the COPD or IPF were classified.

B: The ROC curves for application of scoring coefficients of 15 principal components to categorize the samples as IPF or COPD. On the x–axis is plotted false positive rate (1-specificity) and y–axis is sensitivity plotted. The area under curve is corresponding to the numerator of the Mann-Whitney U-statistic matching the scoring coefficients disseminations between IPF and COPD. The red Line at 45 degree represents 50% classification of IPF vs COPD participants.

2.4 DISCUSSION

Here we have demonstrated that using advance bioinformatics modeling and computation, we were able to in distinguishing different types of diseases such as IPF or COPD from control. The differences between the two devastating diseases (IPF and COPD) could be distinguished from each other by using microarrays mRNA transcripts expression datasets. We were able to reduce 29232 probes by using significant p value, Benjamini and Hochberg multiple correction and critical cutoff level of 1.5-fold change for IPF vs control, IPF vs COPD. The ROC of IPF vs control was significant with an AUC = 0.90 (95% CI), and ROC of IPF vs COPD was significant with an AUC= 0.88 (95% CI).

In this study, 44 transcripts distinguished IPF from control and are involved in several pathways such as cell–cell signaling, cytokine activities, response to wounding and extracellular matrix proteins accumulation. These processes are consistent with activation of TGFB1 in pulmonary fibrosis. TGFB1 stimulates fibroblasts to play major roles in the fibrotic foci and is a well-known driver of fibrotic lung remodeling. It is possible that increased mRNA in the lung transcripts may serve as an indicator of active fibrotic lung remodeling.

Transcripts expression has been previous used for cancer diagnosis and classification of non-Hodgkin (61). Also the transcripts expression of prostate cancer tissues and the neighboring prostate cancer areas have been studied (62). The breast cancer biomarkers (63), ovarian cancer biomarker (Ca -125) (64, 65) and colon cancer biomarker (CAE) have been widely studied at stages of the diseases and used in clinical setting (66, 67). However, in IPF, there is no single clinical biomarker for diagnosis, prediction, prognostication or for predicting drug response.

The limitation of our study is that the IPF subjects or COPD subjects were sampled without categorizing them in sub-phenotypes of diseases. Nonetheless, we used a larger dataset that is capable to reduce the subject to subject variations.

Taken together, the estimate of transcripts expression can distinguish IPF from COPD or control. The comparison of transcripts expression between IPF vs COPD was possible first by using t-test, logistic regression model (59) and followed by specific tests to predict probability and classification of the diseases. Therefore, future studies to explore these methods to distinguish stages of disease or group of disease should be encouraged.

3.0 PAPER 2: CROSS-SECTIONAL AND SERIAL PERIPHERAL BLOOD PROTEIN PROFILES ARE PREDICTORS OF OUTCOMES IN IDIOPATHIC PULMONARY FIBROSIS

3.1 INTRODUCTION

Idiopathic pulmonary fibrosis is a chronic and lethal lung disease (11) mainly affecting the elderly (50 to 79 years old) and mainly males (68). To date, lung transplantation remains to be the most beneficial therapy for IPF. The clinical scenarios of IPF are variable and the outcomes of the disease are unpredictable and there is no single reliable biomarker that has been established for IPF (5, 6). Physiological and functional measurements such as six minute walk distance (6MWD) and pulmonary function test are not yet accepted as surrogate biomarkers for IPF mortality (69, 70). However, those established measurement such as FVC, and DLCO are used to observe disease progression in clinical setting (71).

In the recent years, many investigators have examined transcripts or proteins that could be potential biomarkers of IPF. For example, increased MMP7 transcript and protein in serum, plasma and lung tissues have been noted in IPF patients and have been proposed as possible predictors of IPF mortality (13, 16, 30). Similarly, circulating fibrocytes in peripheral blood, (29), increased Chemokine (C-C motif) ligand 18 (CCL18) (72), and mucin1, cell surface associated (MUC1 aka KL-6) (20) in IPF serum and C-X-C motif chemokine ligand 13

(CXCL13) (Chemokine Motif) in IPF plasma (21) have been advocated as predictors of prognosis of IPF disease. Interestingly a polymorphism of MUC5B was shown to have association with subgroup of IPF subject (73). Despite these persuasive studies, the need of reliable and reproducible peripheral blood biomarkers of IPF that support discovery of pathogenesis or therapeutic target of IPF remains un-met.

In this study we replicated ICAM1, VCAM1, S100A12, IL-8 and SFTPD proteins as biomarkers for IPF in cross-sectional study; and MMP10, ICAM1, VCAM1, S100A12, IL8 and TIMP1 proteins as predictors of mortality in longitudinal study. The combined personal clinical and molecular mortality prediction index (PCMI) of the protein biomarkers improves the probability of predicting IPF mortality. We published portions of these results as an abstract in American thoracic society conference book in May 2014 (74).

3.2 METHODS

3.2.1 Searchlight Protein Array

All patients were evaluated at the University of Pittsburgh Medical Center (UPMC) and studies were approved by the Institutional Review Board (IRB) at the University of Pittsburgh. The diagnosis of IPF was established on the basis of American Thoracic Society (ATS) and European Respiratory Society (ERS) criteria (11). The recruitment and data collection (75), and 61 subjects were recruited from ongoing study: “the Genomic and Proteomic Analysis of Disease Progression in Idiopathic Pulmonary Fibrosis (GAP)”. The repeated blood samples and clinical

assessments were collected in intervals of 3-4 months. The 24 proteins in table 1 were assessed in 250 samples by using Searchlight Protein Array (Aushon Biosystems, Billerica, MA).

Table 3-1: Candidate proteins for outcome in Idiopathic Pulmonary Fibrosis detected in peripheral blood.

Gene Symbol	Proteins	Ref. in IPF
IL1-alpha	Interleukin 1, alpha	(76)
IL2	Interleukin 2	(77)
IL6	Interleukin 6	(78, 79)
IL8	Interleukin 8	(80)
IL10	Interleukin 10	(81)
IFN -alpha	Interferon, alpha	(82)
ICAM1	Intercellular adhesion molecule 1	(30, 83)
VCAM1	Vascular cell adhesion molecule 1	(30, 84)
MPO	Myeloperoxidase	(85)
MIP (AQP0)	Major intrinsic protein of lens fiber	(85)
MMP1	Matrix metallopeptidase 1	(13)
MMP3	Matrix metallopeptidase 3	(86)
MMP7	Matrix metallopeptidase 7	(30)
MMP8	Matrix metallopeptidase 8	(87)
MMP9	Matrix metallopeptidase 9	(7)
MMP10	Matrix metallopeptidase 10	(88)
TIMP1	TIMP metallopeptidase inhibitor 1	(65)
SFTPA1	Surfactant protein A1	(89)
SFTPD	Surfactant protein D	(89)
S100A12	S100 calcium binding protein A12	(30)
AGER	Receptor for advanced glycation End products	(16)
LEP	Leptin	(90)
TGFB1	Transforming growth factor, beta 1.	(91)
SPP1	Secreted phosphoprotein 1	(92)

3.2.2 **Microarray**

We extracted total RNA from 63 Control lungs, 154 COPD lungs, and 83 IPF lung tissues obtained from the Lung Tissues Research Consortium (LTRC). The experimental materials,

procedures, Samples collection, IRB, and statistical analysis have been described previously (84). The data is available at http://www.nhlbi.nih.gov/research/resources/ltrc.

3.2.3 Protein Isolation and Western Blot Analysis

Lysates from control, IPF, or COPD lungs were harvested following the manufacturers' protocol (Thermo Fisher Scientific™, Rockford, IL). The concentrations of protein were measured (Bicinchoninic acid assay, Pierce, Rockford, IL). For Western blot analysis, equal amounts of cellular extracts (10µg) were separated on 8% SDS-PAGE gels and transferred to PVDF-Plus membranes (GE Osmonics, Trevose, PA). Western blots analysis were performed by using antibodies against MMP10 (1:1000; Novo biological Littleton, CO), GAPDH (1:500; Novo biological Littleton, CO). After incubation with the respective secondary antibodies, specific bands were visualized by autoradiography using enhanced chemiluminescence according to the manufacturer's instructions (PerkinElmer Life Sciences, Boston, MA). Densitometry was performed using the shareware, ImageJ (http://rsbweb.nih.gov/ij/).

3.2.4 Immunohistochemistry

Paraffin embedded IPF lungs, COPD lungs and control lungs were obtained from University of Pittsburgh Health Sciences Tissue Bank (Pittsburgh, PA). Tissue slides were deparaffinized in serials ethanol and rehydrated in PBS. Slides were blocked for one hour in 10% BSA and incubated with primary MMP10 antibody overnight. A biotinylated donkey anti-rabbit secondary antibody was incubated for one hour after three times washes with PBS. The

experimental details were previously described (93). The images were visualized with Olympus™ microscope, PROVIS (Olympus America Inc., Melville, NY).

3.2.5 Statistical Analysis of Data

Values were presented as mean or median ± SD. Group comparisons were made using an unpaired, two-tailed Student's t-test for normally distributed data. A level of $p < 0.05$ was considered statistically significant. Survival analysis of the cross sectional study was performed using Cox Time–independent model to examine covariates that change over time. Here baseline samples are used and the subjects were followed to event, or lost subjects follow up, or transplantation occurred. However, Cox Time dependent model examines covariates that change over time and it weighs more for samples collected before the events. Here serial samples are used and the subjects were followed to event, or lost subjects follow up, or transplantation occurred. In the biomarkers values were log transformed and dichotomized at the quartiles of the protein values. The p values were estimated by log rank test and the significance was established at 5%. Kaplan Meier curves plotted for each protein biomarker showing two lines: lower 3^{nd} quartile and higher 3^{th} quartile. Personal clinical and molecular mortality prediction index (PCMI) was created by using the following formula; PCMI = 114 *I(Male) + 2*(100%-FVC% predicted) + 3*(100% - DLco %predicted) + 111*I(Biomarker value ≥ median). Given male, I =1, female, I=0, Biomarker value ≥ median =1, and Biomarker value ≤ median =0 and Kaplan Meier survival curve was plotted using dichotomized PCMI level at median. The ROC of average PCMI of the individual significant biomarkers was estimated by using C statistic to

achieve optimal AUC. We applied Cox Time-dependent model to examine covariates that change over time in the longitudinal cohort and plotted their Hazards ratios (95% CI) while adjusted or unadjusted for sex and age.

3.3 RESULTS

3.3.1 Demographic and clinical information of subjects with IPF

In this study, male consist of 73% of the cohort, former smokers were 72%, never smokers were 24% and current smokers 4%. Of this population 62% died or had lung transplant within 3 years, among them 82 % were males. To replicate the association of gender to IPF disease, we estimated the impact of gender in our cohort and found that male gender is 146 % associated with IPF disease ($p= 0.02$). Although it is well established that IPF disease is related to age, in this study the youngest IPF subject was 53 yr old and the oldest IPF subject was 79 yr. old. The overall average of age of the subjects in this cohort was 70 ± 6 yr. The pulmonary function test revealed that average of DLco = 46±17% years and FVC 64±15% predicted. The mean and median of the each protein measured and the PCMI are described in Table 3.2.

Table 3-2: The mean and median of candidate proteins for outcome in idiopathic pulmonary fibrosis

Proteins	Mean	Median	Std. Dev
ICAM1(pg/ml)	666302.2	618277.3	196097.7
VCAM1(pg/ml)	2304026	2146211	863914.8
ENRAGE(pg/ml)	23843.4	44559.6	85493.2
SFTPD(pg/ml)	219742.6	219742.6	118.7
PCMI	Mean	Median	Std. Dev
ICAM1	371.5	351.6	118.8
VCAM1	373.3	376.1	104.0
S100A12	373.8	386.5	104.9
SFTPD	369.7	364.4	118.7
Combined PCMI	372.0	388.5	101.8

3.3.2 **Cross-sectional protein biomarkers and personal clinical and molecular mortality prediction index (PCMI) predict mortality in IPF subjects.**

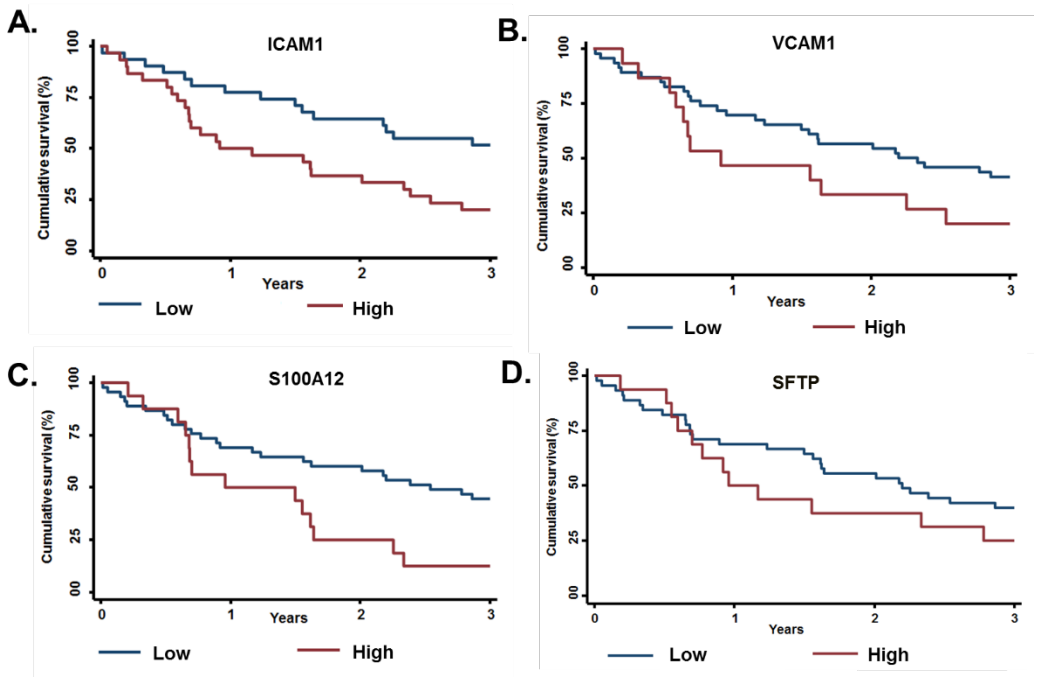


Figure 3-1: Survival analysis of a cross sectional study for candidate biomarkers that predict IPF survival.

Kaplan Meier survival for (A) ICAM1, (B) VCAM1, (C) S100A112, and (D) SFTPD. Blue line ≤3rd quartile, Brown line ≥3rd quartile. In each case $p < 0.05$

In this cross-sectional study, we replicated proteins reported as biomarkers for predicting IPF mortality (13, 30, 94). We measured 250 samples from 61 IPF subjects to determine the level of 24 plasma proteins. Among these the univariate analysis classified 4 proteins, ICAM1 Hazard Ratio (HR) = 3.8 Confidence Interval (CI) (1.3, 11.3) p=0.02, VCAM1 HR=2.7 CI (1.1, 6.5)

p=0.02, S100A12 HR=1.6, CI (1.1, 2.3) p=0.01, and SFTPD HR=1.5, CI (1.1, 2.2), p=0.01) as predictors of IPF mortality after adjusting for age and gender (Figure 3.1).

The personal clinical and molecular mortality prediction index (PCMI) of biomarkers predict the IPF mortality accumulatively and individually as presented in Figure 3.2

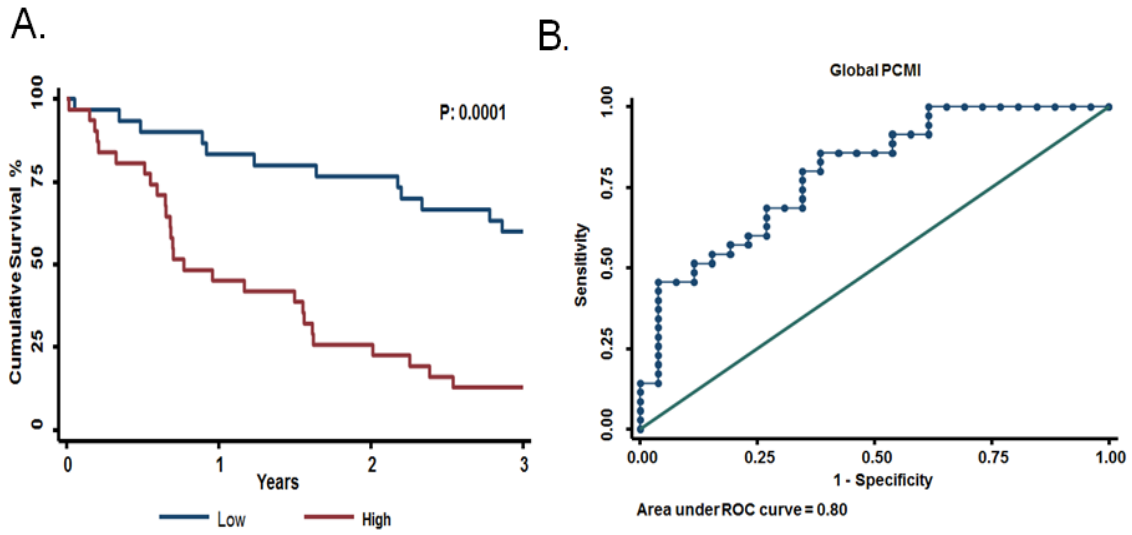


Figure 3-2: Receiver operating characteristic curves for predicting IPF mortality by using personal clinical and molecular mortality Index (PCMI).

(A) PCMI was generated using this formula; PCMI = 114 *I(Male) + 2*(100%-FVC% predicted) + 3*(100% - DLco % predicted) + 111*I (Biomarker value ≥ median). Given male =1, female =0, Biomarker value ≥ median =1, and Biomarker value ≤ median =0. Kaplan Meier plot for survival categorized at the median of PCMI: Blue line stands for PCMI ≤ median and brown line stands for PCMI = ≥ median. (B) ROC of combined individual significant biomarkers based on their PCMI with AUC = 80%, sensitivity 80and specificity 65%).

3.3.3 **Predictors for IPF mortality in longitudinal study**

To identify whether biomarkers in serial observation would better classify peripheral blood protein, we measured the concentrations of 24 proteins in IPF plasma by using multiplex ELISA (Aushon Biosystem, Billerica, MA). We hypothesize those changes in protein expressions could serve as novel predictors of mortality and disease monitoring. Using cox time dependent model, VCAM1, ICAM1, S100A12, MMP10, TIMP1, and IL-8 resulted as predictors of prognosis for IPF mortality while adjusted for gender and age as shown in Figure 3. 3.

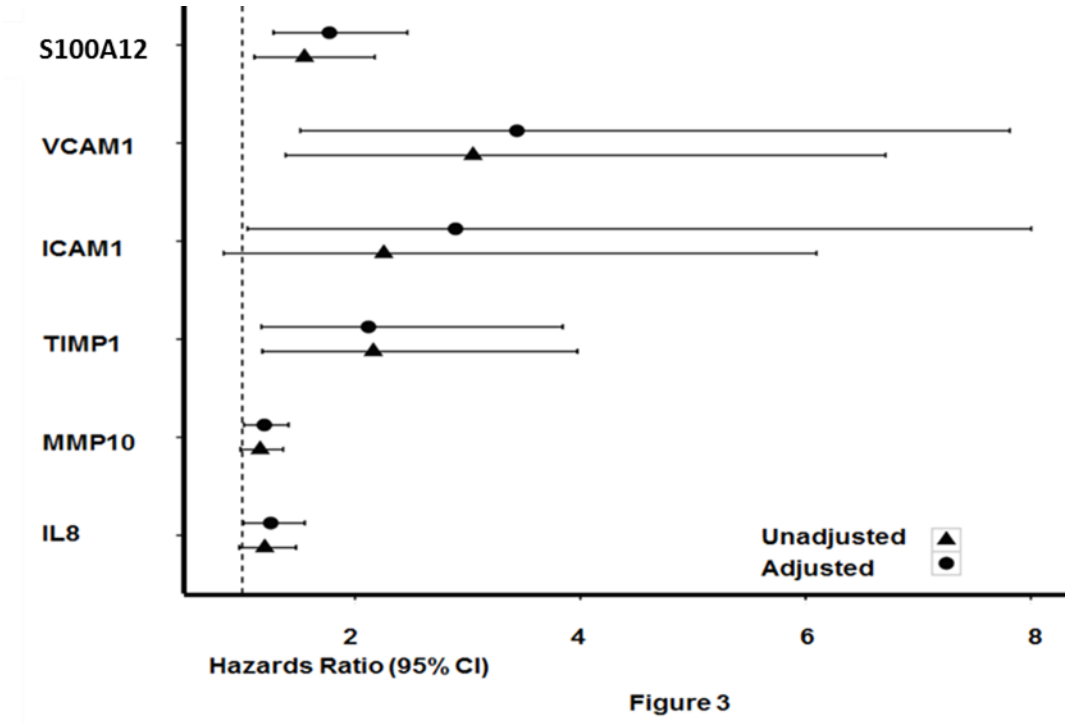


Figure 3-3: Prediction of IPF mortality by using Cox Time dependent model in serial study.

Cox Time – dependent model was used to examine covariates that change over time. In the figures are Hazards ratios (95% CI) of adjusted (Black round plots) or unadjusted (Triangle black) for sex and age. The Hazards ratios were plotted: S100A12 (HR =1.59, P=0.012), VCAM1 (HR = 2.71, P= 0.024), ICAM1 (HR=3.83, P= 0.015), TIMP1 (HR=2.12, P = 0.013), MMP10 (HR=1.19, P = 0.035) and IL8 (HR= 1.25, P = 0.042).

3.3.4 **MMP10 gene and protein expression were higher in IPF compared to COPD and control.**

We compared the transcripts expression level of MMP10 in control, IPF and COPD lungs and we found increased MMP10 in IPF lungs as shown in Figure 3.4A. To compare MMP10 protein

levels in IPF lungs to control lungs and COPD lungs, we performed western blot analysis and we found increased pro MMP10 and active MMP10 protein in IPF lungs (Figure 3.4B-D). To localize MMP10 in IPF lungs, we performed Immunohistochemistry (IHC) analysis. In normal histology lungs, MMP10 was located in the airway epithelialium region (Figure 3.5A). In IPF lungs MMP10 is in dense fibrotic region (Figure 3.5B). To Compare COPD as a disease control, we performed IHC in COPD lungs and found less MMP10 in COPD lungs compared to IPF lungs (Figure 5C).

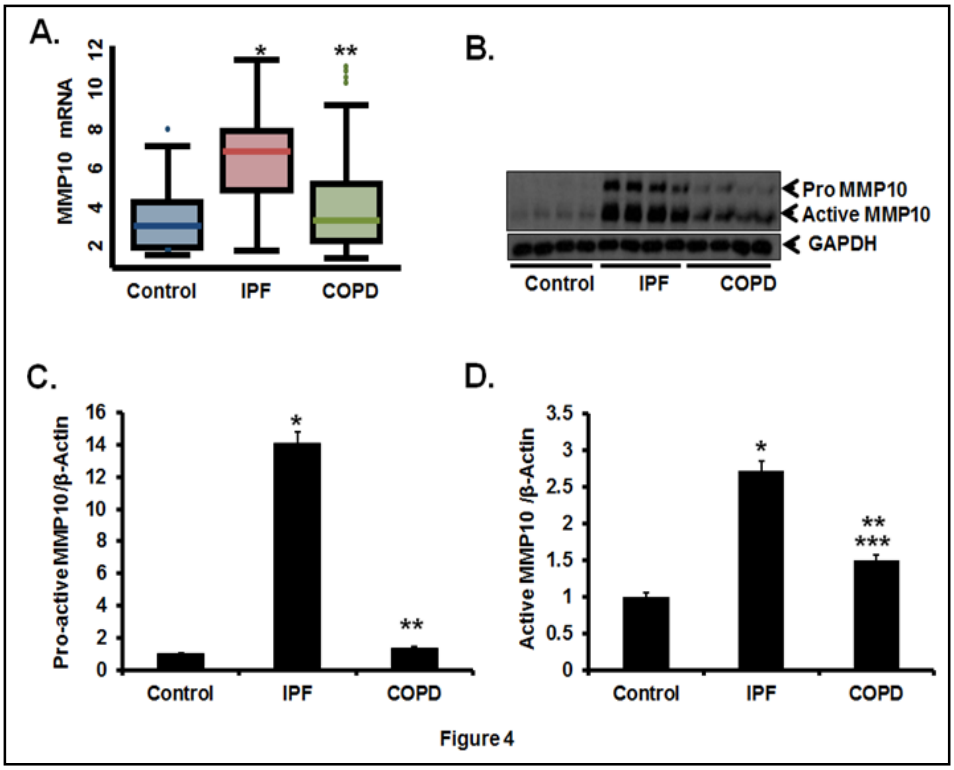


Figure 3-4: MMP10 distinguishes IPF from control and COPD.

(A) MMP10 mRNA level increased in IPF lung tissue (*n*= 83) compared to Control lung (*n*=63) and COPD lungs (*n*= 155) in microarrays data analysis, t-test p=<0.05 (B) MMP10 protein level in IPF, control and COPD. (C-D) demonstrates the densitometry of pro MMP10 and active MMP10. P<0.05 for t-test: control vs IPF (*) and IPF vs COPD (**), control vs COPD (***).

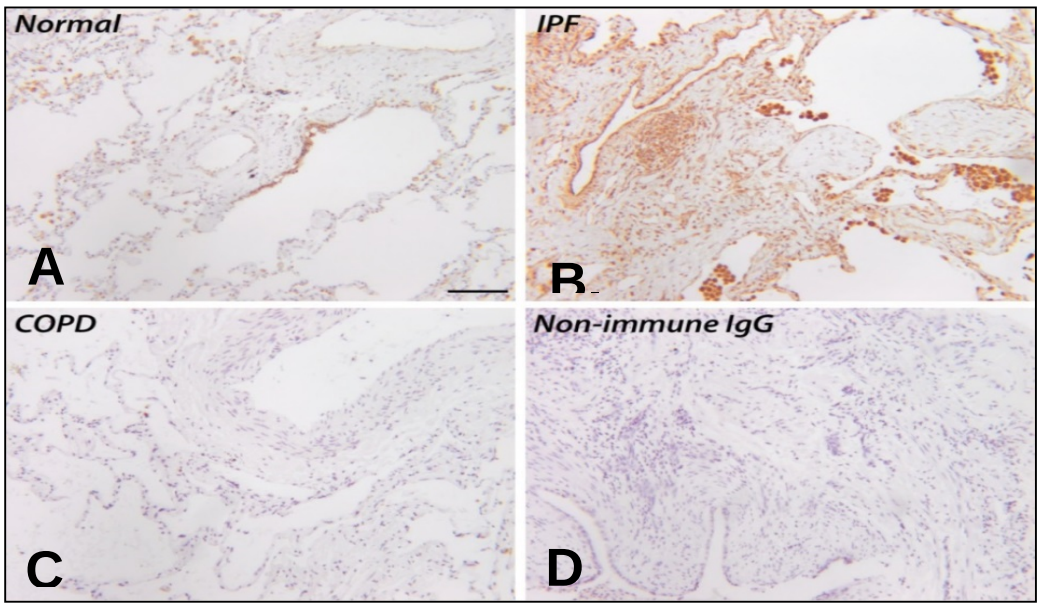


Figure 3-5 Localization of MMP10 in the control, IPF and COPD lungs.

MMP10 protein in the lung sections with brown color stained with DAB and counterstained by using Haematoxylin. (A) Control lung, (B) MMP10 localized mostly in the IPF lungs in the fibroblastic foci, macrophages and lymphocytes. Less MMP10 was expressed in (C) COPD and (D) negative control.

3.4 DISCUSSION

In this study we demonstrated circulating protein biomarkers are predictors of IPF mortality in cross-sectional and longitudinal analysis. These results confirmed previous reports (30) and added additional protein biomarkers with potential clinical and research implication. In cross sectional study, we found ICAM1, VCAM1, S100A12, and SFTPD are predictors of IPF

mortality. Additionally, the personal clinical and molecular mortality prediction index (PCMI) of the biomarkers demonstrated predictive probability of IPF mortality. In longitudinal study, the analyses revealed S100A12, VCAM1, ICAM1, TIMP1, MMP10, and IL-8 were predictors of IPF mortality. Our data confirmed that in both cross-sectional and longitudinal studies VCAM1, ICAM1 IL-8 and S100A12 were consistent predictors of IPF mortality while MMP10 and TIMP1 were only novel predictors for IPF mortality in longitudinal study. Therefore, stratification of IPF subjects using serial study maybe vital since the biomarkers status precedes the outcome events.

In this study we found MMP10 and TIMP1 transcripts were associated with increased mortality in IPF. The finding MMP10 mRNA and protein are increased in IPF lungs as well as in peripheral blood maybe important to understand the mechanism of IPF pathology. MMP10 has a role in cancer metastasis (95) and is increased in human epithelial cancers (96). We and other investigators have reported that MMP10 predicts IPF mortality (74, 88) and is increased in lung and plasma of IPF (88). Interestingly, we identified MMP10 immunostaining in epithelial cells, leukocytes and macrophages and fibrotic areas of the lung. This is consistent with the high rate of lung cancer in IPF subjects (97).

Tissue inhibitor metalloproteinase 1 (TIMP1) is a glycoprotein expressed in most human cells and tissues. As an inhibitor of MMP degradation, TIMP1 regulates extracellular matrix components (98, 99) and in tissues remodeling, repair, embryonal development and angiogenesis (98). Although the role of TIMP1 in IPF diseases is unknown, it seems that both the MMPs inhibitory activities and and MMPs indepenent activity are involved in IPF disease.

Among the other transcripts, we replicated ICAM1, VCAM1, S100A12, IL-8. In addition, SFTPD was also observed to be increased in IPF subject, which has been found to be

increased in acute lung injury patients when compared to control subjects. SFTPD predicts mortality in IPF subjects (30). Although the mechanism of SFTPD movement into blood is unclear, it is possible it occurs through the following mechanisms: (1) hypersecretion of SFTPD by diffuse hyperplasia of type II cells, (2) increased basolateral release and deposition into the extracellular matrix, or (3) increased lung epithelial and endothelial cell injury permitting leakage from pulmonary vasculature leading to increased serum/plasma level.

ICAM1 is present in low concentrations on the surface of leukocytes and endothelial cells and ICAM1 facilitates leukocyte transmigration by binding to integrin subunit beta 2 (ITGB2 aka leukocyte function associated Antigen-1), macrophage adhesion ligand-1 (MAC-1), and fibrinogen (100, 101). In this regard, transforming growth factor beta 1, an integral mediator of pulmonary fibrosis, displays cross-talk with VCAM1 raising a potential signaling mechanism that may be relevant in human IPF (102). Indeed, our observation that elevated VCAM1 concentrations in plasma of IPF subjects are linked with mortality offers support to premise that VCAM1 is involved in pathogenesis of IPF and may be considered as a biomarker. More studies are needed to determine the biological role of VCAM1 in pulmonary endothelium to elucidate potential molecular mechanism(s) of VCAM1 up-regulation and implicate this adhesion molecule as causally related to the pathobiology.

S100A12 is a protein localized to the cytoplasm and nucleus of cells and involved in the regulation of a number of cellular processes including cell cycle progression and differentiation. Serum S100A12 levels increase in patients that develop acute lung injury compared to those who do not develop ALI following postoperative sepsis (103). IL-8 is a chemoattractant and angiogenesis factor that signals inflammatory cells to site of inflammation (104) and we found the high level of IL-8 predict mortality.

Although we have very impressive results, the following limitations in studying peripheral blood proteins as biomarkers are evident: (1) many proteins interact and bind with each other leading to false positive results especially in long duration of storage, (2) processing samples right from blood draw, freezing or thawing cycles until to target proteins measurement are not globally standardized, and (3) multiplex ELISA prototypes may not be optimized for each protein in an assay of multiple proteins.

In summary, we demonstrated that plasma protein expression levels of S100A12, VCAM1, ICAM1, TIMP1, MMP10, and IL-8 are predictors of IPF mortality over 3 years. This is the first time group of IPF biomarkers have been studied in a longitudinal study. Our data demonstrates that peripheral blood biomarkers have capabilities to predict IPF mortality that could impact the treatment and clinical management of this incurable disease.

4.0 PAPER 3: E2F8 IS A CRUCIAL REPRESSOR OF COLLAGEN1 SYNTHESIS AND DEPLETED BY AN ORPHAN E3 LIGASE

4.1 INTRODUCTION

Idiopathic pulmonary fibrosis (IPF) is a lethal and refractory member of the idiopathic interstitial pneumonias, with a median survival time of ~3 years regardless of treatment (105). IPF is histologically characterized by patchy areas of active fibroblast proliferation (fibroblastic foci) that are associated with minimal inflammation and interspersed with areas of advanced fibrotic cystic changes (honeycombing) and normal appearing lung (11). Lung transplant, while definitive, is therapy for only a fraction of IPF patients, although newer therapeutics may limit loss of pulmonary function (106-108). Studies using transcripts targeting in mice, transcriptomic studies, and better classification of interstitial lung diseases have led to a shift in the understanding of IPF(36). IPF has shifted from "inflammation leading invariably to fibrosis" to a more complex model that includes repeated epithelial injury, and aberrant epithelial-mesenchymal crosstalk leading to fibroblast activation, proliferation, and ultimately scarring (109). Roles for new pathways within coagulation, apoptosis, angiogenesis, and development have been proposed(110). Regardless of the proximate events, the final common pathway is the unremitting accumulation of fibroblasts and the deposition of extracellular matrix (36). In IPF,

specifically, collagens 1A1 and 5A1 and elastin are excessively deposited in the interstitial compared to healthy subjects (31) (32).

Extracellular matrix components cross-talk with fibroblasts affecting their behavior (111) and thus understanding how these components are produced is critical in IPF. The abundance of type 1 collagens (COL1A1 and COL1A2) regulated by mRNA synthesis and stability; for example, transforming growth factor beta 1 (TGFB1) and interleukin 4 (IL4) transcriptionally activate the *COL1A1* and *COL1A2* genes in human lung fibroblasts through combinatorial transactivation by SP1, NFKB, and STAT6 (112). TGFB1 also increases COL1A1 mRNA stability (84). Other transcriptional partners such as Yin Yang 1 (113), cis-acting promoter elements (CCAAT box) (114), and upstream signals (Rac1/MLK3/JNK/AP-1) activat*e COL1A1* gene transcription (115). Very few repressors of *COL1A1,* such as Fra-1 or FOXF1 have been identified although definitive regulation of the *COL1A1* promoter by these proteins has not been assessed (116, 117).

E2 transcriptional factor 8 (E2F8) is a member of E2F transcription factors that are divided into two groups, transcriptional activators (E2F1- E2F3) and transcriptional repressors (E2F4-E2F8). E2F family of transcription factors is valuable in mammalian in regulating cell cycles. E2F8 functions as repressor and has two DNA-binding domains (DBDs) that resemble the DBDs of typical E2F as well as their dimerization partner (DP) protein. E2F8 binds a DNA sequences (15-base pair DNA fragment) similar to typical E2F proteins (118). Interestingly, E2F8 was found to block fibroblast cells proliferation and silencing E2F8 induces cell proliferation (119, 120). E2Fs engage other factors such as Rb or DP proteins to control gene transcription (33). E2F1 deficient mice have more biliary fibrosis through Egr-1 (a zinc finger protein) (34), but overexpression of E2F1 inhibits type II collagen formation in cartilage (35).

E2F8 regulates embryonic development and the cell cycle by inhibiting E2F1, and controls angiogenesis and lymphoangiogenesis in zebrafish (121). High-level expression of E2F8 regulates cell proliferation and tumorigenesis in lung cancer and hepatocellular carcinoma (122-124). E2F8 facilitates hepatic fatty liver by regulating FABP3, that may ultimately be associated with hepatic fibrosis (125), suggesting that its function is highly cell- and context-specific.

Ubiquitination tags proteins for degradation, either by the proteasome or lysosome, to regulate diverse processes (126),(127). Ubiquitin linked to a target protein is orchestrated by a series of enzymatic reactions involving an E1 ubiquitin-activating enzyme, ubiquitin transfer from an E1-activating enzyme to an E2-conjugating enzyme, and last, generation of an isopeptide bond between the substrate's ε-amino lysine and the C-terminus of ubiquitin catalyzed by an E3-ubiquitin ligase (128). Of the E3s, the SCF superfamily is the best studied (129). The SCF complex has a multi-unit core consisting of Skp1, Cullin1, and Rbx1 (130). This complex also contains a receptor subunit, termed F-box protein, which recruits substrates to the E3 catalytic core (131, 132). Three families of F-box proteins exist: the L family contains a leucin-rich repeat (LRR) motif; the W family contains a WD repeat motif; and the O family F-box protein contains unknown motifs (133). The protein level of E2F8 is decreased in IPF lungs, fibroblasts and bleomycin induced mice lung fibrosis. It was targeted for ubiquitination and proteasomal degradation in hLF mediated by SCF F-box protein, FBX016 that leading to increased COL1A1 synthesis. hLF transfected with Myc-tagged E2F8 plasmids encoding E2F8 suppressed fibroblasts proliferation or COL1A1 synthesis and E2F8 shRNA constructs induced fibroblasts proliferation or COL1A1 synthesis. Interestingly, hLF co-transfected with a *GLucCol1A1* promoter-reporter construct with plasmids encoding E2F8 suppressed and induced by E2F8

shRNA. Collectively, these studies offer a foundation for further research on the mechanistic roles of E2F8 and FBX016 in IPF pathogenesis.

4.2 METHODS

4.2.1 Materials.

Microarray reagents, arrays and scanner were purchased from Agilent Technologies (Wilmington, DE). The following antibodies were used: Collagen 1A1 (Rockland, Limerick, PA), β –actin (Sigma-Aldrich, St. Louis, MO), MMP10 (Novous Biological, Littleton, CO), E2F8 (Abcam, Cambridge, MA), FBX016 and Myc (Santa Cruz Biotechnology, Santa Cruz, CA). The following reagents used: cycloheximide (Inducer of protein degradation), MG132 (Inhibitor of proteasomal protein degradation), Leupeptin (Inhibitor of lysosomal protein degradation), IgG, and ubiquitin (Santa Cruz Biotechnology, Santa Cruz, CA), bleomycin and trichrome stain kit (Sigma-Aldrich, St. Louis, MO), the miRNA mini Kit and Effecten transfection reagent (Qiagen, Louisville, KY), the primers for collagen1A1 and qRT-PCR (ABI, Foster City, CA), E2F8 shRNA (Dharmacon, Lafayette, CO), the *GLucCol1A* promoter reporter construct (GeneCopoeia, Rockville, MD), the CytoSelect BrdU cell proliferation ELISA kit (Cell Biolabs, San Diego, CA) and Western Lightning Plus ECL (PerkinElmer, Boston, MA).

4.2.2 **Microarrays**

Lung Tissue Resource Consortium (LTRC) comprised of 4 clinical centers with core laboratories of tissues, radiology and clinical data that were coordinated in collecting data of COPD and IPF. For this study, we used 109 controls and 134 IPF samples. All samples were well phenotyped clinically, radiographically and physiologically and the diagnoses of IPF were established on the basis of American Thoracic Society (ATS) and European Respiratory Society (ERS) Criteria (11). Total RNA were extracted, labeled and hybridized to Agilent 44k whole human genome microarrays (Agilent Technologies, Wilmington, DE).

4.2.3 **Lung Protein isolation and western blot analysis**

Lysates from control and IPF lungs were collected (Thermo Fisher Scientific™, Rockford, IL), protein concentrations were measured (Pierce's Bicinchoninic acid, Pierce, Rockford, IL). For Western blot analysis, equal amounts of cellular extracts (10µg) were separated on 8-10 % SDS-PAGE gels and transferred to PVDF-Plus membranes (GE Osmonics, Trevose, PA). Western blots were performed with antibodies against E2F8, MMP10, FBX016, Anti-Collagen type 1 and Anti-Myc were diluted (1:1000) and β-actin (1:10000). After incubation with the respective secondary antibodies, specific bands were visualized by autoradiography using enhanced chemiluminescence according to the manufacturer's instructions (PerkinElmer Life Sciences, Boston, MA). Densitometry was performed using the shareware, ImageJ (http://rsbweb.nih.gov/ij/).

4.2.4 **Cells culture, RNA extraction, and real-time PCR experiments.**

Early passages (1-3) of primary human lung fibroblasts (hLF) from the University of Pittsburgh Tissue Bank were cultured in F-12 medium supplemented with 10% fetal bovine serum (FBS) and antibiotic-antimycotic (Invitrogen, Carlsbad, CA). All experiments were performed on cells at 70-80% confluence. Treatment with E2F8 plasmid or a control empty vector and control shRNA or E2F8 shRNA were performed for 24 hrs. Immunoblots were performed as described above. The cellular total RNA were extracted from cultured cells on plates by QIAol lysis reagent and purified using the miRNeasy mini Kit from Qiagen (Louisville, KY). The RT-PCR was performed according to the manufacture's protocol as described previously (134).

The proliferation was detected using a CytoSelect BrdU cell proliferation ELISA. To determine BrdU incorporation into cellular DNA during cell proliferation, cells were grown on a 96-well cell culture plate and transfected with E2F8 plasmid or a control empty vector and control shRNA or E2F8 shRNA using Effecten transfection reagent. After 12 hrs, medium was changed and cells were incubated for an additional 24 hrs. An aliquot of BrdU was then added to the medium and cells were incubated for an additional 3 hrs at 37^0C. After washing with PBS, cells were fixed for 30 min and BrdU incorporation into total cellular DNA was determined using anti-BrdU antibody.

4.2.5 **Bleomycin murine model of lung fibrosis.**

All animals' procedures were executed in accordance with approval of University of Pittsburgh Institutional Animal Care and Use Committee. Male and female C57BL/6 mice (6 to 8 weeks

old) were deeply anesthetized and bleomycin at 3 U/kg or saline control was administered intratracheally in a volume of 50 μL. After 16 days mice were sacrificed (CO2 inhalation and cervical dislocation) and the lungs were excised for determination of Collagen content using Masson trichrome staining. The images were visualized (PROVIS Olympus America Inc., Melville, NY).

4.2.6 Statistical Analysis

The group comparisons between diseased and control subjects were performed using an unpaired two-tailed Student's *t*-test for normally distributed data. A level of $p<0.05$ was considered statistically significant. The bands on immunoblots were quantified by using ImageJ, and then densitometric ratios were calculated. The ratios data were statistically analyzed to discriminate their differences by using an unpaired student *t*-test. Microarrays data were analyzed by cyclic-LOESS normalization (135), Genomica (136), and SAM (136). A 5% false discovery rate was used for significance and expressed as absolute fold change.

4.3 RESULTS

4.3.1 E2F8 protein is a target for ubiquitin-mediated degradation in human lung fibroblasts (hLF).

In mRNA screens of transcription factors in IPF lungs versus control human lungs, we assayed levels of various members of the E2F family and first observed a ~ 3-fold increase in steady-state

mRNA for E2F8 (Fig. 1A). Interestingly, when assaying levels of E2F8 protein decreased in IPF (*n*=5) compared to controls (n=5) (Fig. 2 A-B). Low levels of E2F8 were also seen in the fibroblasts isolated from IPF (n=4) lungs versus control (n=4) (Fig. 2C-D). To understand if E2f8 is rapidly degraded in IPF fibroblasts through proteasome, we exposed to hLF cycloheximide (CHX) to assess protein stability and observed that E2F8 has a short half-life (t ½ ~1-2 h) (Fig. 3A) consistent with other transcription factors. E2F8 degradation was inhibited by a proteasome inhibitor (MG132) but not a lysosomal inhibitor (leupeptin). Transcription of a plasmid encoding Hemagglutinin tagged-ubiquitin (HA-Ub) in hLF increased E2F8 polyubiquitination (Fig. 3B) and HA-Ub expression was sufficient to trigger E2F8 degradation (Fig. 3C).

We next hypothesized that E2F8 itself must undergo proteolysis in IPF fibroblasts because it is depleted and yet its mRNA is increased 3-fold versus normal lung (Figure 1A). One possibility is that E2F8 is targeted by a ubiquitin E3 ligase for it elimination in IPF. E2F8 is a phosphoprotein and because many E3 SCF-based proteins target phosphorylated substrates for ubiquitination (137), we performed an unbiased screen randomly examining ubiquitin E3 ligase receptor components, called F-box proteins, that might be expressed in IPF lung and trigger E2F8 degradation. We found that FBX010, FB0X015, and FBX016 mRNA were increased IPF lungs compared to control lungs (Fig. 4A). Upon expression of one of these plasmids, FBX016, decreased levels of immunoreactive E2F8 compared to a control plasmid, FBX024 (Fig. 4B). In a bleomycin model of lung fibrosis, E2F8 protein levels were reduced and FBX016 levels modestly increased (Fig. 4C). To confirm specificity of FBX016 targeting E2F8, we performed Co-IP experiments where FBX016 was detected in the E2F8 immunoprecipitates (Fig.4D).

Figure 4-1 E2F8 mRNA in IPF lungs compared to control lungs in fold change.

(A). Analysis of the micro mRNA levels of E2F transcription factor family members expressed in IPF lungs as fold change vs control lungs. The mRNAs were extracted from LGRC data by microarray profiling of lung tissues from control subjects (n=109) and subjects with IPF (n=134).

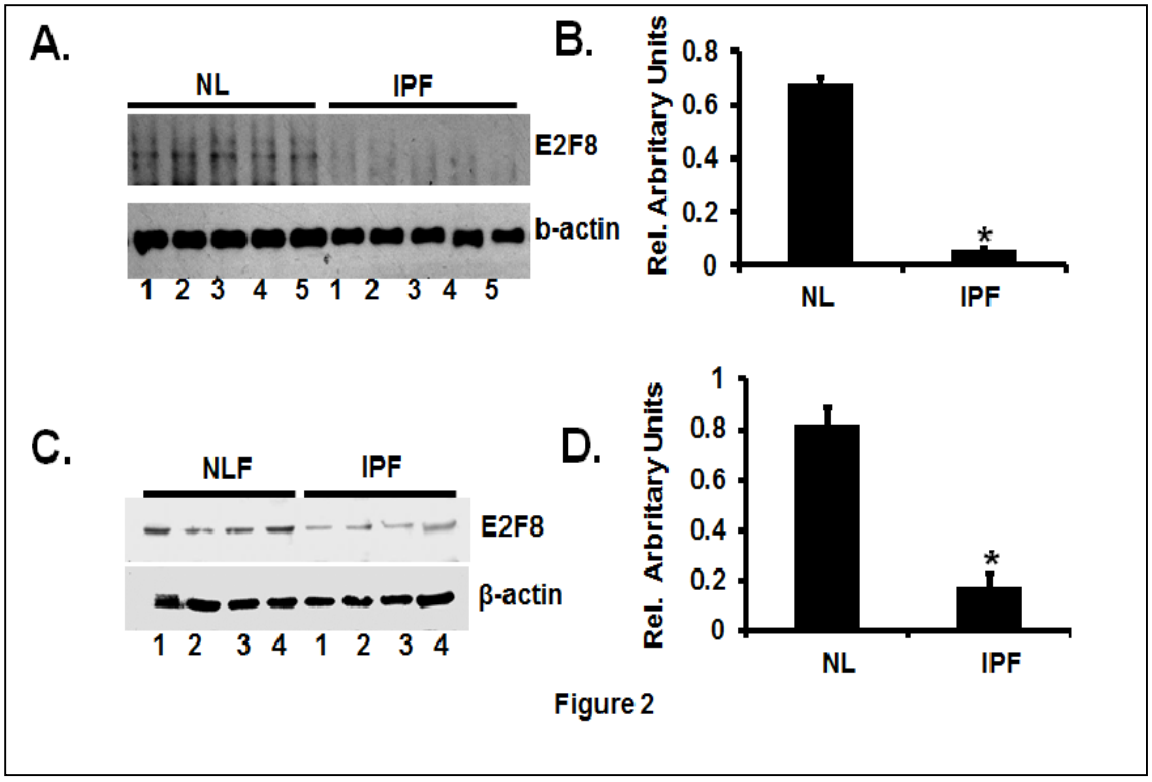


Figure 4-2: E2F8 protein level is decreased IPF lungs and fibroblasts.

E2F8 protein (94 kDa) in control (n=5) lungs and IPF (n=5) lungs. (B) Densitometry quantification of the immunoblot of E2F8 in lung tissues lysate (Image J). (C) Analysis of E2F8 protein levels in fibroblasts isolated from control (n=4) and IPF (n=4) lungs. (D) Densitometry quantification of the immunoblot of E2F8 in fibroblasts (Image J).

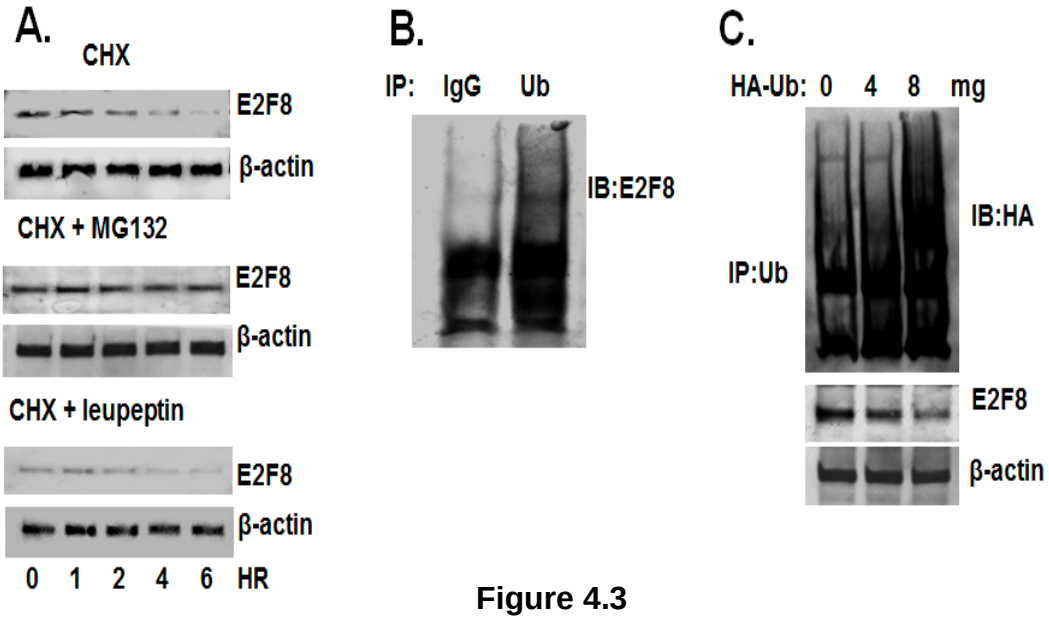


Figure 4.3

Figure 4-3Ubiquitin-proteasome degradation of E2F8 in human lung fibroblasts.

(A) Levels of E2F8 in hLF exposed to of cyclohexamide (CHX, 20 mg/ml) alone or in combinations with MG132 (20 mg/ml) or leupeptin (100mg/ml). (B) Ubiquitination of E2F8. Lysates from lung fibroblasts were subjected to immunoprecipitation with IgG or ubiquitin antibodies followed by immunoblotting for E2F8. (C) Expression of hemagglutinin-tagged ubiquitin (HA-Ub) in hLF induced degradation of E2F8 and increased total ubiquitination of proteins in cell lysates. Human lung fibroblasts transfected with various concentrations of plasmid encoding HA-Ub were analyzed by immunoblotting for E2F8 and β-actin.

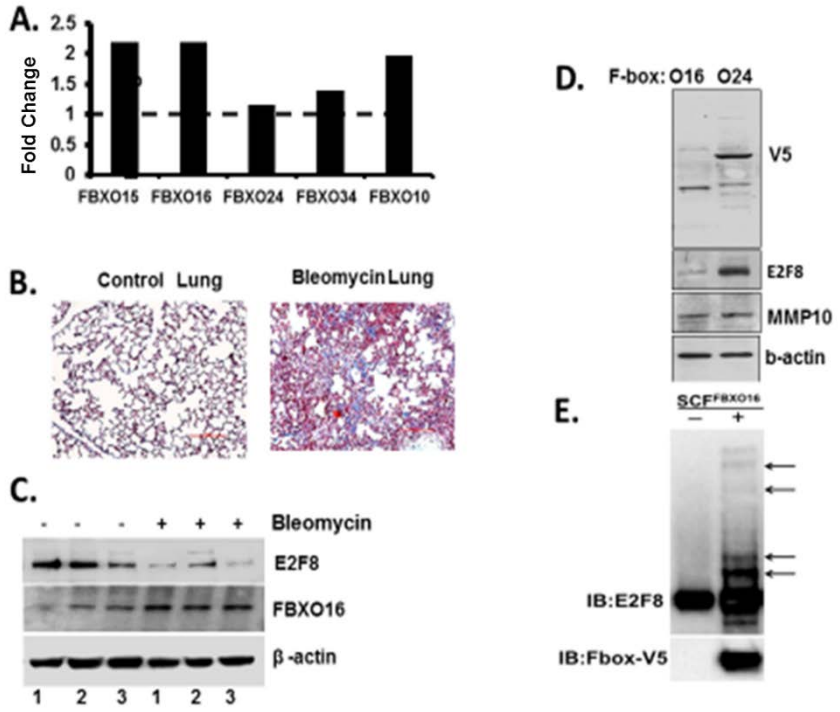


Figure 4-4: Fbxo16 triggers E2F8 degradation in human and mouse fibrotic lung.

Expression of F-box proteins in IPF lungs versus control lungs. The mRNAs were extracted from LGRC data: Microarray profiling of lung tissues from control subjects (n=109) and subjects with IPF (n=134). (B) Histology of control lung and bleomycin induced mice lung fibrosis (C) FBXO16 targets E2F8 for degradation. The hLF were transfected with V5-tagged F-box proteins (0.4 mg) and a control FBX024 plasmid. Cells were than lysed for V5, E2F8, MMP10 and β-actin levels by immunoblotting. MMP10 was used as a negative control. (C) Levels of expression of E2F8 and FBX016 proteins in lysates from a bleomycin- induced mouse model of lung fibrosis were analyzed by immunoblotting. (D). Ubiquitination reaction showing ability of FBX016 to directly polyubiquitinate (arrows) E2F8.

4.3.2 **E2F8 regulates collagen1A1 synthesis in lung fibroblasts.**

We examined if E2F8 transcriptionally suppresses collagen 1 levels in hLF because E2F1 knockout mice are less prone to biliary fibrosis (34) and this protein may regulate type II collagen formation. We ectopically expressed a plasmid encoding Myc-tagged E2F8 in hLF and observed that levels of collagen 1A1 decreased with transfection of increasing amounts of E2F8 plasmid (Fig. 5A). Conversely, depleting cells of E2F8 increased collagen 1A1 mRNA content in cells (Fig. 5B). Further, in these studies we also assayed steady-state collagen 1 (Col 1) mRNA in cells and observed that overexpression of E2F8 decreased Col 1 mRNA levels (Fig. 5D) whereas depletion of the transcription factor increases the transcript levels for this matrix-associated protein (Fig. 5E). In a bleomycin fibrosis model, Col 1 levels increased (Fig. 5C). To confirm effects of E2F8 on inhibition of Col 1 gene transcription, we co-transfected human lung fibroblasts with a *GLucCol1A1* promoter reporter construct together with a plasmid encoding E2F8 or an empty vector (Fig. 6A). Ectopic E2F8 expression in cells strongly repressed promoter-reporter activity over 50% versus control (Fig. 6B). In reverse experiments, we knocked down E2F8 using shRNA in hLF which increased E2F8 silencing increased *GLucCol1A1* promoter reporter activity almost 1.5 fold (Figure 4.6D).

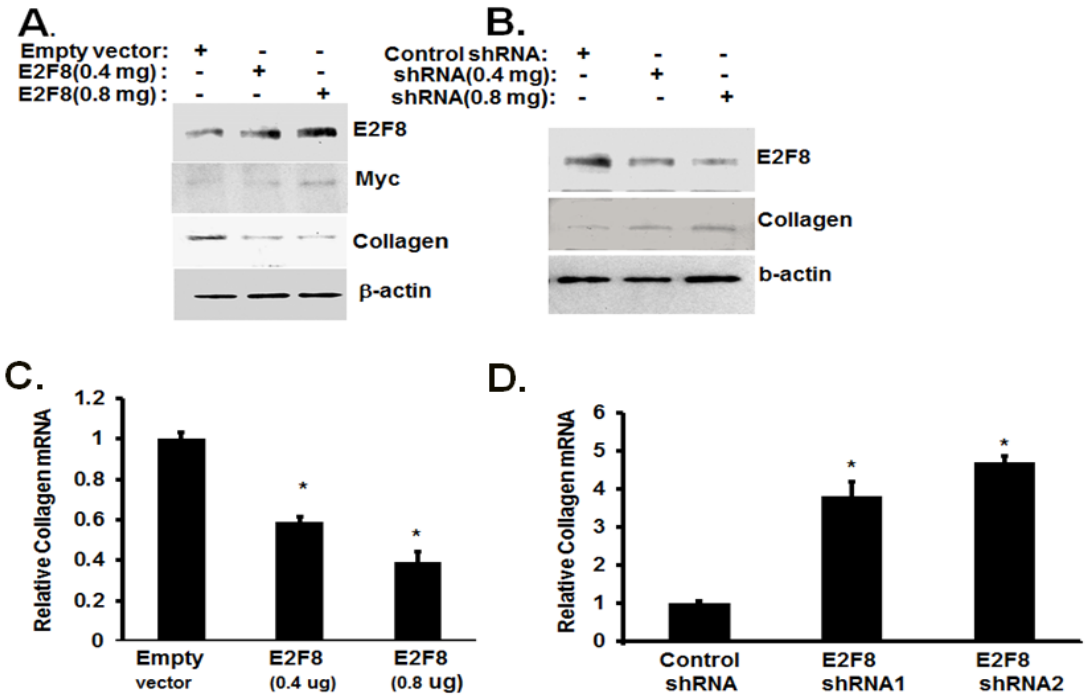


Figure 4-5 E2F8 modulates collagen 1 levels.

hLF were transfected with (A, D) various amounts of Myc-tagged E2F8 plasmids encoding E2F8 or an empty vector, or (B, E) one of two E2F8 shRNA constructs or a control shRNA. (A, B) Cells were then lysed and levels of E2F8 and collagen 1 levels were determined by immunoblot analysis. (C) The levels of collagen 1 expression in lungs of mice treated with bleomycin. Upper: Normal lung structure of lung histology and below is damaged lung histology with blue color showing collagen 1 in a bleomycin- induced fibrosis. (D, E) Total cellular RNA was extracted from the lung fibroblasts using TRIzol reagent. Quantitative real-Time PCR was then performed using the Taqman Gold kit according to the manufacturer's instructions. mRNA of collagen 1 was measured. *P<0.05

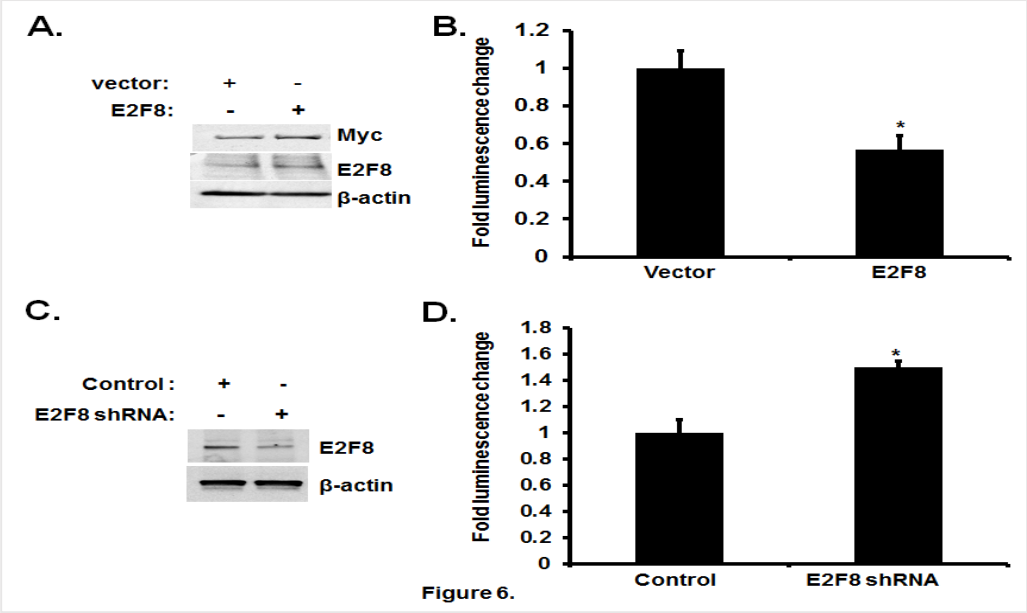


Figure 4-6: E2F8 represses collagen 1 gene transcription.

Normal Lung fibroblasts were co-transfected with *a GLucCol1A* promoter-reporter construct and plasmids encoding E2F8 or an empty vector (A-B), or an E2F8 shRNA construct or a control shRNA (C-D). After 48 h cell culture, medium was collected for GLuc luminescent assays using a Secrete-Pair™ Gaussia Luciferase Assay Kit which detects naturally secreted GLuc following the manufacturer's instructions. Cell lysates were then analyzed by immunoblotting to determine transfection efficiencies (A, C). The data are from n=3 separate experiments. *P<0.05.

4.3.3 **E2F8 plasmid inhibits and depletion of E2F8 induced lung fibroblast proliferation**

We hypothesized that E2F8 might play a role in IPF fibroblasts proliferation and the formation characteristic fibroblastic foci. Thus, we examined if E2F8 modulates fibroblast growth

behavior. We ectopically expressed Myc-tagged E2F8 in normal hLF (NHLF) or in hLF isolated from IPF lungs and assayed cell proliferation using BrdU labeling. We also depleted this transcription factor using shRNA as described above. As shown in figure 7, transfection of E2F8 plasmid modestly reduced cell proliferation by ~25-35% in both NHLF and IPF fibroblasts (Fig. 7A, C). In contrast, depleting cells of E2F8 increased cell growth despite low levels of E2F8 in IPF cells (Figure 4.7B, D).

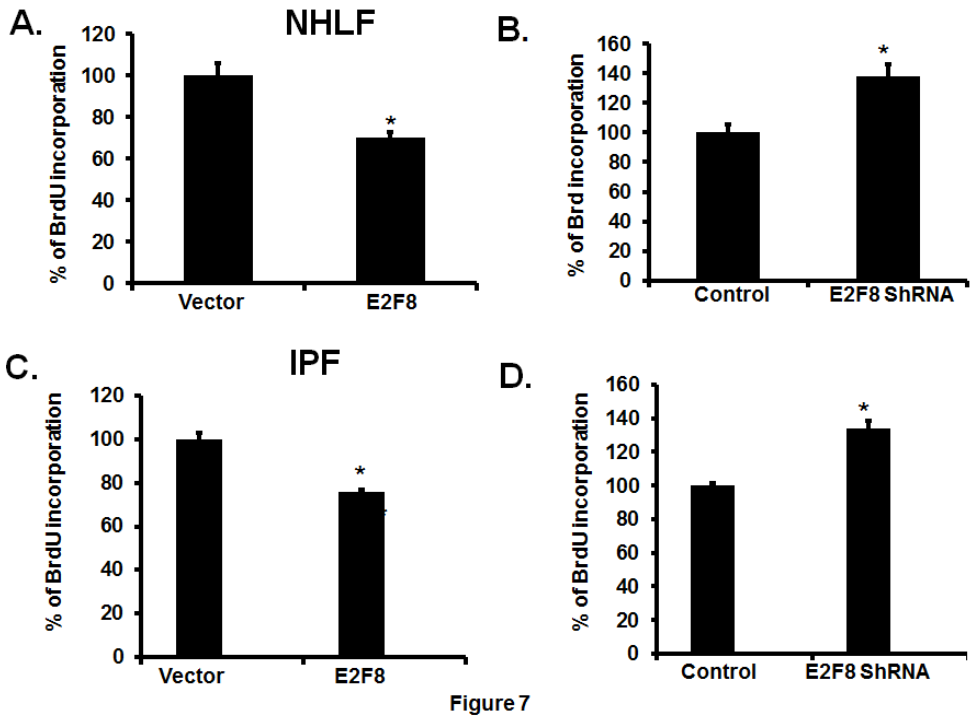


Figure 4-7 E2F8 modulates fibroblast proliferation.

The proliferation of hLF was detected using a CytoSelect^{TM}BrdU cell proliferation ELISA kit. To determine BrdU incorporation into cellular DNA during cell proliferation, normal hLF (NHF) (A, B) or IPF fibroblasts (C, D) on a 96-well cell culture plate were transfected with: (A, C) plasmids encoding E2F8 or an empty vector, or (B, D) E2F8 shRNA or a control shRNA using Effecten transfection reagent. The BrdU incorporated into cellular DNA during cell proliferation was detected using an anti-BrdU antibody. The data from each panel represents n=4 separate experiments. *$P<0.05$.

4.4 DISCUSSION

In this study, we investigated the role of E2F8 in the pathogenesis of IPF disease and the results suggest that: (A) E2F8 mRNA is increased in IPF lung, (B) E2F8 is a crucial transcriptional repressor of collagen 1A1 levels in human lung fibroblasts, (C) E2F8 protein levels are depleted in IPF lung and IPF lung fibroblasts because of SCF-FBX016-mediated ubiquitin-proteasomal degradation, (D) loss of E2F8 appears to contribute at least in part to the pathobiology of fibroblastic growth and fibrotic response observed in humans with IPF disease. This data offers a biologic foundation for future investigations of the molecular regulation of this transcriptional repressor (E2F8) and collagen1A1 synthesis in IPF.

Previous studies have found that E2F8 levels increase in lung cancer (122), hepatocellular carcinoma (138) (123) and melanoma (139). However the biological functions appears to cell type and context specific. In cancer cells, silencing E2F8 inhibited lung cancer cell proliferation, and tumor growth. Increased E2F8 predicted mortality of adenocarcinoma subjects (122). Taken together E2F8 has opposite functions in cancer cells compared to the reports in normal mouse embryonic fibroblasts. E2F8 inhibits cells proliferation and silencing E2F8 induces cell proliferation (119, 120). Since in this study we examined E2F8 in normal human lung fibroblasts and IPF lung fibroblasts, our data replicated that suppression of fibroblasts proliferation by overexpression of E2F8 and knocking down E2F8 in lung fibroblasts yielded to induction of fibroblasts proliferation. Although transcript expression of E2F8 is increased in IPF as in cancers, we found that E2F8 protein is decreased in IPF lung fibroblasts. To understand these phenomena, we investigated whether a FBX016 could limit E2F8, by

ubiquitination and proteasomal degradation in the fibroblasts of IPF subjects. FBX016 might serve as a unique molecular target for therapeutic intervention in fibrotic illness using F-box pharmaceutical inhibitors.

Because the histology of usual interstitial pneumonia exhibits active fibroblasts proliferation (fibroblastic foci), honeycombing and accumulation of extracellular matrix such as COL1A1, fibronectin and elastin (31, 32, 36), we co-transfected E2F8 plasmid in hLF with a *GLucCol1A* promoter reporter construct and noted that ectopic E2F8 expression in cells repressed COL1A1 promoter-reporter activity. Complementary knock down of E2F8 in hLF increased *GLucCol1A* promoter reporter activity. This data suggest that E2F8 is a key inhibitor of COL1A1 synthesis in human lung fibroblasts.

TGFB1 is a key mediator of fibrotic lung remodeling and it is conceivable that decreased expression of E2F8 protein or increased E2F8 mRNA concentrations in IPF lungs or fibroblasts may function as a sign of active fibrotic lung transformation because low level of E2F8 in mouse lungs was observed in a murine model of bleomycin-induced pulmonary fibrosis. As such the possibility that bleomycin is involved in regulation of E2F8 is relevant for physiological process of lung fibrosis. We hypothesized that E2F8 could play a role in IPF fibroblasts proliferation and forming characteristic foci in IPF lung. Thus, we examined if E2F8 modulates fibroblast growth behavior. We ectopically expressed Myc-tagged E2F8 plasmid in normal hLF or in hLF isolated from IPF lungs and assayed cell proliferation using BrdU labeling. We also depleted this transcription factor using shRNA. Transfection of E2F8 plasmid modestly reduced cell proliferation by ~25-35% in both NHLF and IPF fibroblasts. Similarly, depleting cells of E2F8 increased cell growth despite low levels of E2F8 in IPF cells. Our studies also raise additional questions as to the role of transcriptionally activated E2F8 on modulating fibroblast phenotypic

behavior because E2F8 has been reported to be involved in modulating growth activity in mouse and human uterine decidualization (140), and placental development (141).

Although we offer interesting results, we are aware of limitation that we did not performed *in vivo* mouse model of E2F8 gene loss or gain experiment. However, our findings *in vitro* experiments on human lung fibroblasts are relevant mechanistic information and should encourage future IPF studies.

Our data suggest that E2F8 transcript is increased and E2F8 protein is decreased by proteasomally degradation in IPF lung and this was confirmed by treating fibroblasts from IPF subjects with the proteasome inhibitor, MG132 and that these E2F8 protein levels can be restored by interrupting the mechanisms that mediate protein degradation. These results are novel and important because they demonstrate that E2F8 may act to repress COL1A1 gene transcription and this could lead to a better understanding of IPF pathobiology and potential therapeutic targets.

5.0 CONCLUSION

In this project we investigated the complex problems of IPF disease: (1) lack straight forward method of diagnosis, (2) IPF is a chronic, unpredictable, and incurable lung disease, (3) IPF is still named “Idiopathic” that means the etiology of this particular usual interstitial pneumonia is unknown. Therefore, to address these 3 fundamental issues, we applied the following experimental methodologies: transcriptomic analysis, plasma protein biomarkers discovery, molecular mechanistic evaluations.

We demonstrated that advance bioinformatics modeling and computation, we could distinguish IPF or COPD from control through their specific transcript expressions. We were able to reduce 29232 transcripts to 44 transcripts. The ROC of IPF vs control was significant with an AUC = 0.85 (95% CI), sensitivity 85% and specificity 91%. The ROC of IPF vs COPD was significant with an AUC= 0.87 (95% CI), sensitivity 86% and specificity 90%. Although this is very interesting results, we are aware that utilitizing lungs transcripts has major limitations in clinical practice because biopsy is invasive procedure and microarrays is an expensive and labor intensive method. The next step that could bring this bioinformatics finding to clinical settings is to measure the corresponding proteins of the transcripts in the blood samples. Following that experiment, a diagnostic protein biomarkers panel could be created after a validation study is being performed using independent cohort. Therefore, future studies that

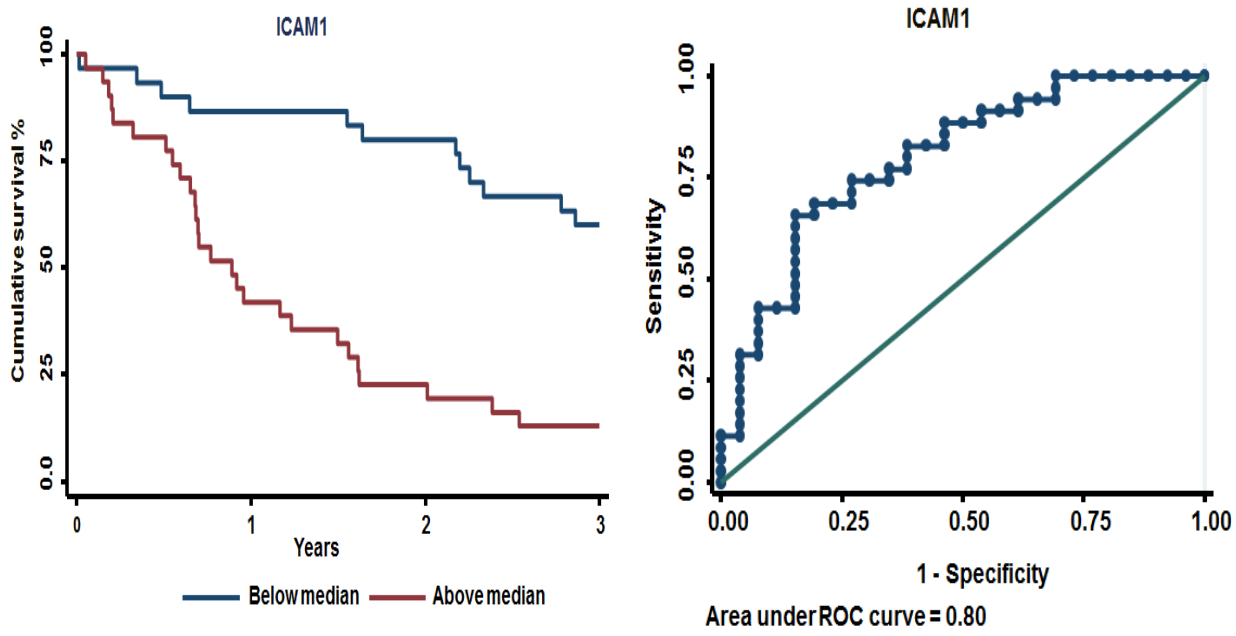
would like to explore these methods to distinguished stages of disease or group of disease should be encouraged.

In the prognostic biomarker study, we replicated previously described predictors of IPF mortality in cross-sectional study. In longitudinal study, our data confirmed that in both cross-sectional and longitudinal studies S100A12, VCAM1, ICAM1, and IL-8 were consistent predictors of IPF mortality and added MMP10 and TIMP1 as novel predictors for IPF mortality in longitudinal study. Therefore, stratification of IPF subjects using serial studies biomarkers will be vital since the concentration of biomarkers status precedes the outcome events weighted more in the analysis. This is the first time group of IPF biomarkers have been studied in a longitudinal study. As IPF is clinically variable disease, stratifying IPF subjects based on risk of mortality could assist in selecting IPF subject for lung transplant and clinical trial studies. Our data demonstrates that peripheral blood biomarkers have capabilities to predict IPF mortality. Clearly, these prognostic biomarkers, along with other reported IPF biomarkers, will require further validation. Nonetheless, this approach appears to be worthy of additional investigation and could impact the treatment and clinical management of this incurable disease

E2F8 is a crucial transcriptional repressor of COL1A1 levels in human lung fibroblasts. Loss of E2F8 appears to contribute, at least in part to the pathobiology of fibroblastic growth and fibrotic response observed in humans with IPF disease. These new biochemical and molecular observations offers a biologic foundation for future projects aiming to investigate the molecular regulation of E2F8, COL1A1 synthesis, and lung fibroblast proliferation in IPF. In addition, these observations support restoration of E2F8 function, possibly by small molecule compounds, as a valid target for IPF therapy.

In summary, our transcriptomic and plasma proteins analyses revealed the utility of biomarker discovery in advancing precision medicine in IPF. Together these finding provide novel insights that may improve IPF diagnosis, prognostic and potential drug discovery.

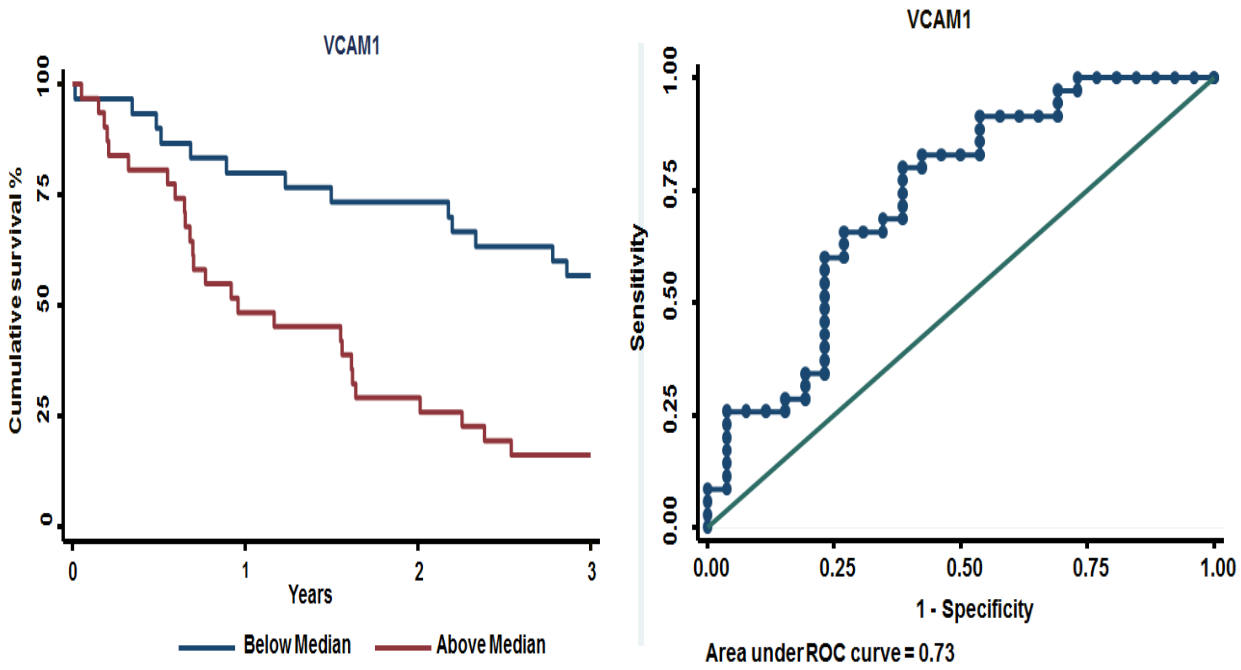
APPENDIX



Supplemental Figure 1: ICAM1.

Receiver operating characteristic curves for predicting IPF mortality by using personal clinical and molecular mortality Index (PCMI). (Left panel) PCMI was generated using this formula; PCMI = 114 *I(Male) + 2*(100%-FVC% predicted) + 3*(100% - DLco %predicted) + 111*I (Biomarker value ≥ median). Given male =1, female =0, Biomarker value ≥ median =1, and Biomarker value ≤ median =0. Kaplan Meier plot for survival categorized at the Median of PCMI: Blue line stands for PCMI ≤ Median and brown line stands for PCMI ≥ Median. (Right

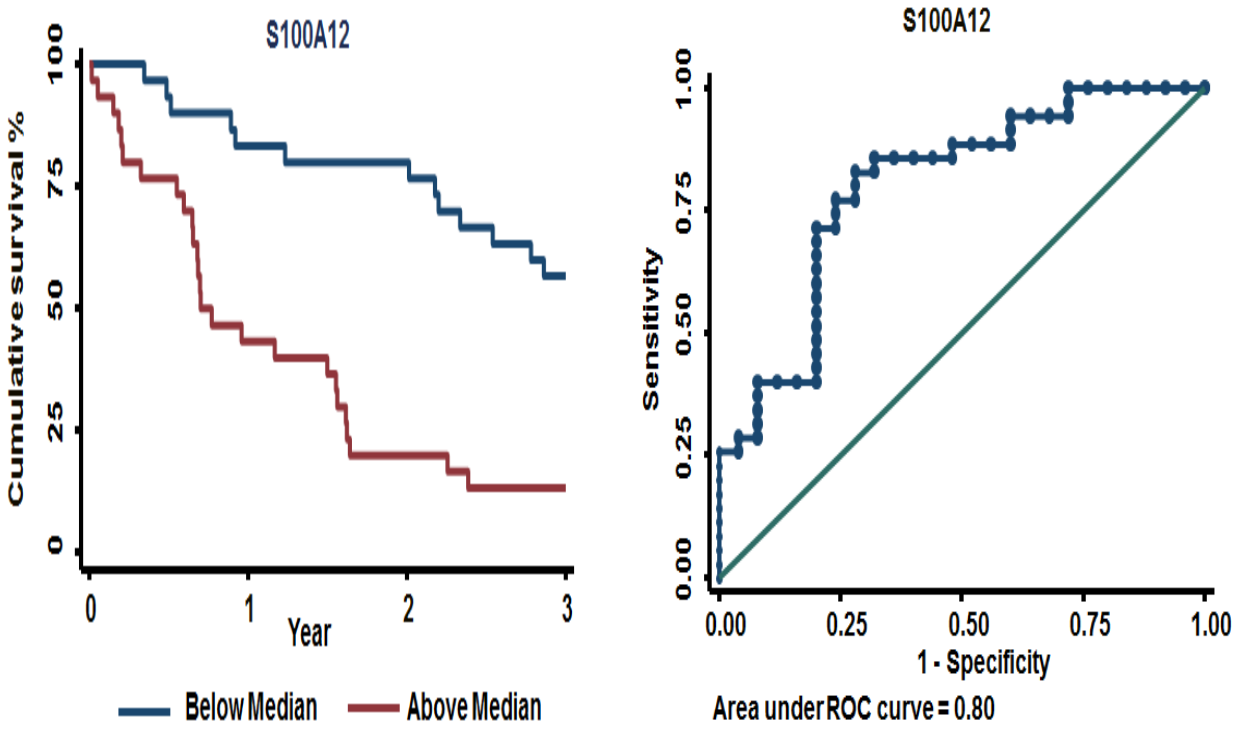
Panel) ROC of combined individual significant biomarkers based on their PCMI with AUC = 80%



Supplemental Figure 2: VCAM1.

Receiver operating characteristic curves for predicting IPF mortality by using personal clinical and molecular mortality Index (PCMI). (Left Panel) PCMI was generated using this formula; PCMI = 114 *I(Male) + 2*(100%-FVC% predicted) + 3*(100% - DLco %predicted) + 111*I(Biomarker value ≥ median). Given male =1, female =0, Biomarker value ≥ median =1, and

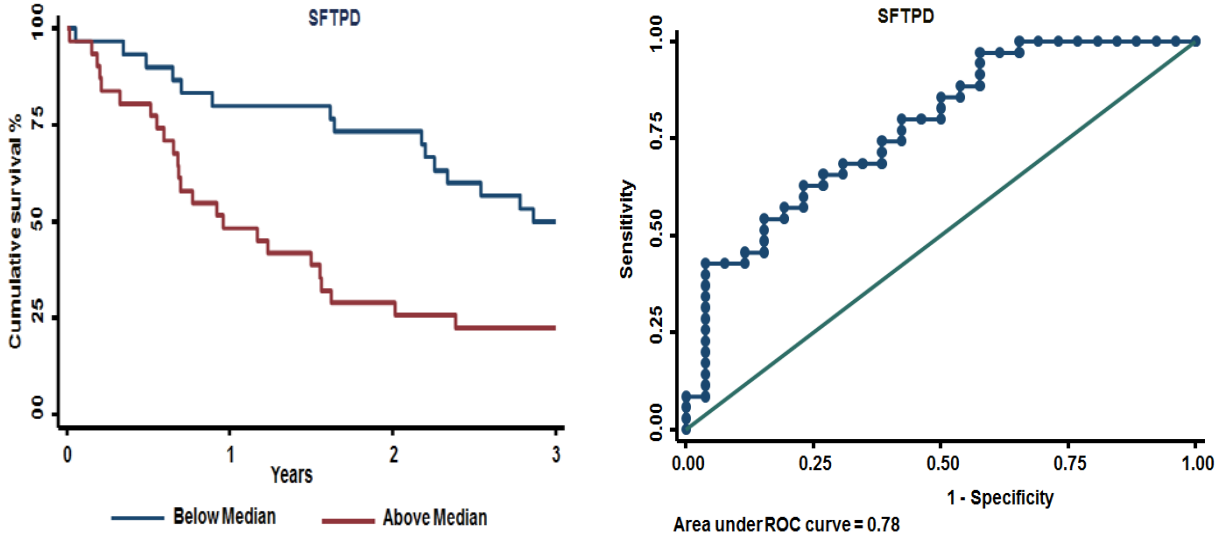
Biomarker value ≤ median =0. Kaplan Meier plot for survival categorized at the Median of PCMI: Blue line stands for PCMI ≤ Median and brown line stands for PCMI ≥ Median. (Right Panel) ROC of combined individual significant biomarkers based on their PCMI with AUC = 73%



Supplemental Figure 3: S100A12.

Receiver operating characteristic curves for predicting IPF mortality by using personal clinical and molecular mortality Index (PCMI). (Left Panel) PCMI was generated using this formula; PCMI = 114 *I(Male) + 2*(100%-FVC% predicted) + 3*(100% - DLco %predicted) + 111*I(

Biomarker value ≥ median). Given male =1, female =0, Biomarker value ≥ median =1, and Biomarker value ≤ median =0. Kaplan Meier plot for survival categorized at the Median of PCMI: Blue line stands for PCMI ≤ Median and brown line stands for PCMI ≥ Median. (Right Panel) ROC of combined individual significant biomarkers based on their PCMI with AUC = 80%



Supplemental Figure 4: SFPD.

Receiver operating characteristic curves for predicting IPF mortality by using personal clinical and molecular mortality Index (PCMI). (Left Panel) PCMI was generated using this formula; PCMI = 114 *I(Male) + 2*(100%-FVC% predicted) + 3*(100% - DLco %predicted) + 111*I(Biomarker value ≥ median). Given male =1, female =0, Biomarker value ≥ median =1, and Biomarker value ≤ median =0. Kaplan Meier plot for survival categorized at the Median of PCMI: Blue line stands for PCMI ≤ Median and brown line stands for PCMI ≥ Median. (Right

Panel) ROC of combined individual significant biomarkers based on their PCMI with AUC = 78%

BIBLIOGRAPHY

1. Fernandez Perez ER, Daniels CE, Schroeder DR, St Sauver J, Hartman TE, Bartholmai BJ, Yi ES, Ryu JH. Incidence, prevalence, and clinical course of idiopathic pulmonary fibrosis: a population-based study. *Chest* 2010; 137: 129-137.

2. King TE, Jr., Bradford WZ, Castro-Bernardini S, Fagan EA, Glaspole I, Glassberg MK, Gorina E, Hopkins PM, Kardatzke D, Lancaster L, Lederer DJ, Nathan SD, Pereira CA, Sahn SA, Sussman R, Swigris JJ, Noble PW. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *N Engl J Med* 2014; 370: 2083-2092.

3. Richeldi L, du Bois RM, Raghu G, Azuma A, Brown KK, Costabel U, Cottin V, Flaherty KR, Hansell DM, Inoue Y, Kim DS, Kolb M, Nicholson AG, Noble PW, Selman M, Taniguchi H, Brun M, Le Maulf F, Girard M, Stowasser S, Schlenker-Herceg R, Disse B, Collard HR. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *The New England journal of medicine* 2014; 370: 2071-2082.

4. Noelker C, Hampel H, Dodel R. Blood-based protein biomarkers for diagnosis and classification of neurodegenerative diseases: current progress and clinical potential. *Mol Diagn Ther* 2011; 15: 83-102.

5. King TE, Jr., Pardo A, Selman M. Idiopathic pulmonary fibrosis. *Lancet* 2011; 378: 1949-1961.

6. Ley B, Collard HR, King TE, Jr. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine* 2011; 183: 431-440.

7. Fernandez IE, Amarie OV, Mutze K, Konigshoff M, Yildirim AO, Eickelberg O. Systematic phenotyping and correlation of biomarkers with lung function and histology in lung fibrosis. *American journal of physiology Lung cellular and molecular physiology* 2016: ajplung.00183.02015.

8. Kim DS, Collard HR, King TE, Jr. Classification and natural history of the idiopathic interstitial pneumonias. *Proceedings of the American Thoracic Society* 2006; 3: 285-292.

9. Lynch JP, 3rd, Huynh RH, Fishbein MC, Saggar R, Belperio JA, Weigt SS. Idiopathic Pulmonary Fibrosis: Epidemiology, Clinical Features, Prognosis, and Management. *Seminars in respiratory and critical care medicine* 2016; 37: 331-357.

10. Zhang Y, Kaminski N. Biomarkers in idiopathic pulmonary fibrosis. *Curr Opin Pulm Med* 2012; 18: 441-446.

11. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, Colby TV, Cordier JF, Flaherty KR, Lasky JA, Lynch DA, Ryu JH, Swigris JJ, Wells AU, Ancochea J, Bouros D, Carvalho C, Costabel U, Ebina M, Hansell DM, Johkoh T, Kim DS, King TE, Jr., Kondoh Y, Myers J, Muller NL, Nicholson AG, Richeldi L, Selman M, Dudden RF, Griss BS, Protzko SL, Schunemann HJ. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *American journal of respiratory and critical care medicine* 2011; 183: 788-824.

12. Katzenstein AL, Myers JL. Idiopathic pulmonary fibrosis: clinical relevance of pathologic classification. *American journal of respiratory and critical care medicine* 1998; 157: 1301-1315.

13. Rosas IO, Richards TJ, Konishi K, Zhang Y, Gibson K, Lokshin AE, Lindell KO, Cisneros J, Macdonald SD, Pardo A, Sciurba F, Dauber J, Selman M, Gochuico BR, Kaminski N. MMP1 and MMP7 as potential peripheral blood biomarkers in idiopathic pulmonary fibrosis. *PLoS medicine* 2008; 5: e93.

14. Kakugawa T, Yokota S, Ishimatsu Y, Hayashi T, Nakashima S, Hara S, Sakamoto N, Kubota H, Mine M, Matsuoka Y, Mukae H, Nagata K, Kohno S. Serum heat shock protein 47 levels are elevated in acute exacerbation of idiopathic pulmonary fibrosis. *Cell stress & chaperones* 2013; 18: 581-590.

15. Hamai K, Iwamoto H, Ishikawa N, Horimasu Y, Masuda T, Miyamoto S, Nakashima T, Ohshimo S, Fujitaka K, Hamada H, Hattori N, Kohno N. Comparative Study of Circulating MMP-7, CCL18, KL-6, SP-A, and SP-D as Disease Markers of Idiopathic Pulmonary Fibrosis. *Disease markers* 2016; 2016: 4759040.

16. Konishi K, Gibson KF, Lindell KO, Richards TJ, Zhang Y, Dhir R, Bisceglia M, Gilbert S, Yousem SA, Song JW, Kim DS, Kaminski N. Gene expression profiles of acute exacerbations of idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine* 2009; 180: 167-175.

17. Prasse A, Stahl M, Schulz G, Kayser G, Wang L, Ask K, Yalcintepe J, Kirschbaum A, Bargagli E, Zissel G, Kolb M, Muller-Quernheim J, Weiss JM, Renkl AC. Essential role of osteopontin in smoking-related interstitial lung diseases. *The American journal of pathology* 2009; 174: 1683-1691.

18. Ishikawa N, Hattori N, Yokoyama A, Kohno N. Utility of KL-6/MUC1 in the clinical management of interstitial lung diseases. *Respiratory investigation* 2012; 50: 3-13.
19. Jaffar J, Unger S, Corte TJ, Keller M, Wolters PJ, Richeldi L, Cerri S, Prele CM, Hansbro PM, Argraves WS, Oliver RA, Oliver BG, Black JL, Burgess JK. Fibulin-1 predicts disease progression in patients with idiopathic pulmonary fibrosis. *Chest* 2014; 146: 1055-1063.

20. Yokoyama A, Kondo K, Nakajima M, Matsushima T, Takahashi T, Nishimura M, Bando M, Sugiyama Y, Totani Y, Ishizaki T, Ichiyasu H, Suga M, Hamada H, Kohno N. Prognostic value of circulating KL-6 in idiopathic pulmonary fibrosis. *Respirology (Carlton, Vic)* 2006; 11: 164-168.

21. Vuga LJ, Tedrow JR, Pandit KV, Tan J, Kass DJ, Xue J, Chandra D, Leader JK, Gibson KF, Kaminski N, Sciurba FC, Duncan SR. C-X-C motif chemokine 13 (CXCL13) is a prognostic biomarker of idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine* 2014; 189: 966-974.

22. Peljto AL, Zhang Y, Fingerlin TE, Ma SF, Garcia JG, Richards TJ, Silveira LJ, Lindell KO, Steele MP, Loyd JE, Gibson KF, Seibold MA, Brown KK, Talbert JL, Markin C, Kossen K, Seiwert SD, Murphy E, Noth I, Schwarz MI, Kaminski N, Schwartz DA. Association between the MUC5B promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *Jama* 2013; 309: 2232-2239.

23. Chien JW, Richards TJ, Gibson KF, Zhang Y, Lindell KO, Shao L, Lyman SK, Adamkewicz JI, Smith V, Kaminski N, O'Riordan T. Serum lysyl oxidase-like 2 levels and idiopathic pulmonary fibrosis disease progression. *Eur Respir J* 2014; 43: 1430-1438.

24. Tajiri M, Okamoto M, Fujimoto K, Johkoh T, Ono J, Tominaga M, Azuma K, Kawayama T, Ohta S, Izuhara K, Hoshino T. Serum level of periostin can predict long-term outcome of idiopathic pulmonary fibrosis. *Respiratory investigation* 2015; 53: 73-81.

25. Korthagen NM, van Moorsel CH, Barlo NP, Ruven HJ, Kruit A, Heron M, van den Bosch JM, Grutters JC. Serum and BALF YKL-40 levels are predictors of survival in idiopathic pulmonary fibrosis. *Respiratory medicine* 2011; 105: 106-113.

26. Gilani SR, Vuga LJ, Lindell KO, Gibson KF, Xue J, Kaminski N, Valentine VG, Lindsay EK, George MP, Steele C, Duncan SR. CD28 down-regulation on circulating CD4 T-cells is associated with poor prognoses of patients with idiopathic pulmonary fibrosis. *PloS one* 2010; 5: e8959.

27. Kotsianidis I, Nakou E, Bouchliou I, Tzouvelekis A, Spanoudakis E, Steiropoulos P, Sotiriou I, Aidinis V, Margaritis D, Tsatalas C, Bouros D. Global impairment of CD4+CD25+FOXP3+ regulatory T cells in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine* 2009; 179: 1121-1130.

28. Reilkoff RA, Peng H, Murray LA, Peng X, Russell T, Montgomery R, Feghali-Bostwick C, Shaw A, Homer RJ, Gulati M, Mathur A, Elias JA, Herzog EL. Semaphorin 7a+ regulatory T cells are associated with progressive idiopathic pulmonary fibrosis and are implicated in transforming growth factor-beta1-induced pulmonary fibrosis. *American journal of respiratory and critical care medicine* 2013; 187: 180-188.

29. Moeller A, Gilpin SE, Ask K, Cox G, Cook D, Gauldie J, Margetts PJ, Farkas L, Dobranowski J, Boylan C, O'Byrne PM, Strieter RM, Kolb M. Circulating fibrocytes are an indicator of poor prognosis in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine* 2009; 179: 588-594.

30. Richards TJ, Kaminski N, Baribaud F, Flavin S, Brodmerkel C, Horowitz D, Li K, Choi J, Vuga LJ, Lindell KO, Klesen M, Zhang Y, Gibson KF. Peripheral blood proteins predict mortality in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine* 2012; 185: 67-76.

31. Vittal R, Mickler EA, Fisher AJ, Zhang C, Rothhaar K, Gu H, Brown KM, Emtiazdjoo A, Lott JM, Frye SB, Smith GN, Sandusky GE, Cummings OW, Wilkes DS. Type V collagen induced tolerance suppresses collagen deposition, TGF-beta and associated transcripts in pulmonary fibrosis. *PloS one* 2013; 8: e76451.

32. Enomoto N, Suda T, Kono M, Kaida Y, Hashimoto D, Fujisawa T, Inui N, Nakamura Y, Imokawa S, Funai K, Chida K. Amount of elastic fibers predicts prognosis of idiopathic pulmonary fibrosis. *Respiratory medicine* 2013; 107: 1608-1616.

33. Zheng N, Fraenkel E, Pabo CO, Pavletich NP. Structural basis of DNA recognition by the heterodimeric cell cycle transcription factor E2F-DP. *Genes Dev* 1999; 13: 666-674.

34. Zhang Y, Xu N, Xu J, Kong B, Copple B, Guo GL, Wang L. E2F1 is a novel fibrogenic gene that regulates cholestatic liver fibrosis through the Egr-1/SHP/EID1 network. *Hepatology (Baltimore, Md)* 2014; 60: 919-930.

35. Scheijen B, Bronk M, van der Meer T, Bernards R. Constitutive E2F1 overexpression delays endochondral bone formation by inhibiting chondrocyte differentiation. *Mol Cell Biol* 2003; 23: 3656-3668.

36. Selman M, King TE, Pardo A. Idiopathic pulmonary fibrosis: prevailing and evolving hypotheses about its pathogenesis and implications for therapy. *Ann Intern Med* 2001; 134: 136-151.

37. Lynch DA, Godwin JD, Safrin S, Starko KM, Hormel P, Brown KK, Raghu G, King TE, Jr., Bradford WZ, Schwartz DA, Richard Webb W. High-resolution computed tomography in idiopathic pulmonary fibrosis: diagnosis and prognosis. *American journal of respiratory and critical care medicine* 2005; 172: 488-493.

38. Checa M, Hagood JS, Velazquez-Cruz R, Ruiz V, Garcia-De-Alba C, Rangel-Escareno C, Urrea F, Becerril C, Montano M, Garcia-Trejo S, Cisneros Lira J, Aquino-Galvez A, Pardo A, Selman M. Cigarette Smoke Enhances the Expression of Profibrotic Molecules in Alveolar Epithelial Cells. *PloS one* 2016; 11: e0150383.

39. Ushiki A, Yamazaki Y, Hama M, Yasuo M, Hanaoka M, Kubo K. Viral infections in patients with an acute exacerbation of idiopathic interstitial pneumonia. *Respiratory investigation* 2014; 52: 65-70.

40. Molyneaux PL, Maher TM. The role of infection in the pathogenesis of idiopathic pulmonary fibrosis. *European respiratory review : an official journal of the European Respiratory Society* 2013; 22: 376-381.

41. Mannino DM, Homa DM, Akinbami LJ, Ford ES, Redd SC. Chronic obstructive pulmonary disease surveillance--United States, 1971-2000. *Respiratory care* 2002; 47: 1184-1199.

42. Mannino DM, Homa DM, Akinbami LJ, Ford ES, Redd SC. Chronic obstructive pulmonary disease surveillance--United States, 1971-2000. *Morbidity and mortality weekly report Surveillance summaries (Washington, DC : 2002)* 2002; 51: 1-16.

43. Olson AL, Swigris JJ, Lezotte DC, Norris JM, Wilson CG, Brown KK. Mortality from pulmonary fibrosis increased in the United States from 1992 to 2003. *American journal of respiratory and critical care medicine* 2007; 176: 277-284.

44. Baumgartner KB, Samet JM, Stidley CA, Colby TV, Waldron JA. Cigarette smoking: a risk factor for idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine* 1997; 155: 242-248.

45. Reikvam H, Oyan AM, Kalland KH, Hovland R, Hatfield KJ, Bruserud O. Differences in proliferative capacity of primary human acute myelogenous leukaemia cells are associated with altered gene expression profiles and can be used for subclassification of patients. *Cell proliferation* 2013; 46: 554-562.

46. Potti A, Mukherjee S, Petersen R, Dressman HK, Bild A, Koontz J, Kratzke R, Watson MA, Kelley M, Ginsburg GS, West M, Harpole DH, Jr., Nevins JR. A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. *The New England journal of medicine* 2006; 355: 570-580.

47. Tamayo P, Scanfeld D, Ebert BL, Gillette MA, Roberts CW, Mesirov JP. Metagene projection for cross-platform, cross-species characterization of global transcriptional states. *Proceedings of the National Academy of Sciences of the United States of America* 2007; 104: 5959-5964.

48. Cheang MC, van de Rijn M, Nielsen TO. Gene expression profiling of breast cancer. *Annual review of pathology* 2008; 3: 67-97.

49. Herazo-Maya JD, Noth I, Duncan SR, Kim S, Ma SF, Tseng GC, Feingold E, Juan-Guardela BM, Richards TJ, Lussier Y, Huang Y, Vij R, Lindell KO, Xue J, Gibson KF, Shapiro SD, Garcia JG, Kaminski N. Peripheral blood mononuclear cell gene expression profiles predict poor outcome in idiopathic pulmonary fibrosis. *Sci Transl Med* 2013; 5: 205ra136.

50. John T, Brenda J-G, David G, Joseph KL, David AS, Avrum S, Mark WG, John Q, Mick C, Thomas R, Lara C, Steven DS, Louis V, Frank CS, Naftali K. Divergent And Convergent Gene Expression Patterns In Well Phenotyped Patients With COPD And IPF. A95 CUTTING EDGE INSIGHTS TO THE PATHOBIOLOGY OF COPD AND INTERSTITIAL LUNG DISEASE: American Thoracic Society; 2011. p. A2274-A2274.

51. Huang Y, Ma SF, Vij R, Oldham JM, Herazo-Maya J, Broderick SM, Strek ME, White SR, Hogarth DK, Sandbo NK, Lussier YA, Gibson KF, Kaminski N, Garcia JG, Noth I. A functional genomic model for predicting prognosis in idiopathic pulmonary fibrosis. *BMC pulmonary medicine* 2015; 15: 147.

52. Schwender H. http://rss.acs.unt.edu/Rdoc/library/siggenes/html/qvalue.cal.html.

53. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America* 2001; 98: 5116-5121.

54. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J. Bioconductor: open software development for computational biology and bioinformatics. *Genome biology* 2004; 5: R80.

55. http://www.bioconductor.org/.

56. Binyamini Y, Hochberg. Controllin the False Discovery Rate: A practical and powerful approach to multiple testing. *J Royal Statistical Soc* 1995; 57: 289-300

57. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. *Behavioural brain research* 2001; 125: 279-284.

58. Joliffe IT, Morgan BJ. Principal component analysis and exploratory factor analysis. *Statistical methods in medical research* 1992; 1: 69-95.

59. Hosmer DW, Taber S, Lemeshow S. The importance of assessing the fit of logistic regression models: a case study. *American journal of public health* 1991; 81: 1630-1635.

60. Vuga LJ, Tedrow JR, Pandit KV, Tan J, Kass DJ, Xue J, Chandra D, Leader JK, Gibson KF, Kaminski N, Sciurba FC, Duncan SR. C-X-C Motif Chemokine 13 (CXCL13) Is a Prognostic Biomarker of Idiopathic Pulmonary Fibrosis. *American journal of respiratory and critical care medicine* 2014; 189: 966-974.

61. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J, Jr., Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000; 403: 503-511.

62. Yu YP, Landsittel D, Jing L, Nelson J, Ren B, Liu L, McDonald C, Thomas R, Dhir R, Finkelstein S, Michalopoulos G, Becich M, Luo JH. Gene expression alterations in prostate cancer predicting tumor aggression and preceding development of malignancy. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2004; 22: 2790-2799.

63. Chen X, Hu H, He L, Yu X, Liu X, Zhong R, Shu M. A novel subtype classification and risk of breast cancer by histone modification profiling. *Breast cancer research and treatment* 2016.

64. Soletormos G, Duffy MJ, Othman Abu Hassan S, Verheijen RH, Tholander B, Bast RC, Jr., Gaarenstroom KN, Sturgeon CM, Bonfrer JM, Petersen PH, Troonen H, CarloTorre G, Kanty Kulpa J, Tuxen MK, Molina R. Clinical Use of Cancer Biomarkers in Epithelial Ovarian Cancer: Updated Guidelines From the European Group on Tumor Markers. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* 2016; 26: 43-51.

65. Zhang Q, Guo Y, Dong R, Dai R, Zhou M. Suppressor of cytokine signaling 1-modulated metalloproteinases and tissue inhibitor of metalloproteinase in pulmonary fibrosis. *Molecular medicine reports* 2015; 12: 3855-3861.

66. Christensen IJ, Brunner N, Dowell B, Davis G, Nielsen HJ, Newstead G, King D. Plasma TIMP-1 and CEA as Markers for Detection of Primary Colorectal Cancer: A Prospective Validation Study Including Symptomatic and Non-symptomatic Individuals. *Anticancer research* 2015; 35: 4935-4941.

67. Chen J, Raju GS, Jogunoori W, Menon V, Majumdar A, Chen JS, Gi YJ, Jeong YS, Phan L, Belkin M, Gu S, Kundra S, Mistry NA, Zhang J, Su X, Li S, Lin SH, Javle M, McMurray JS, Rahlfs TF, Mishra B, White J, Rashid A, Beauchemin N, Weston BR, Shafi MA, Stroehlein JR, Davila M, Akbani R, Weinstein JN, Wu X, Mishra L. Mutational Profiles Reveal an Aberrant TGF-beta-CEA Regulated Pathway in Colon Adenomas. *PloS one* 2016; 11: e0153933.

68. Raghu G, Chen SY, Hou Q, Yeh WS, Collard HR. Incidence and prevalence of idiopathic pulmonary fibrosis in US adults 18-64 years old. *The European respiratory journal* 2016.

69. du Bois RM, Weycker D, Albera C, Bradford WZ, Costabel U, Kartashov A, Lancaster L, Noble PW, Raghu G, Sahn SA, Szwarcberg J, Thomeer M, Valeyre D, King TE, Jr. Ascertainment of individual risk of mortality for patients with idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine* 2011; 184: 459-466.

70. du Bois RM, Weycker D, Albera C, Bradford WZ, Costabel U, Kartashov A, Lancaster L, Noble PW, Sahn SA, Szwarcberg J, Thomeer M, Valeyre D, King TE, Jr. Six-minute-walk test in idiopathic pulmonary fibrosis: test validation and minimal clinically important difference. *American journal of respiratory and critical care medicine* 2011; 183: 1231-1237.

71. Behr J, Bonella F, Bonnet R, Glaser S, Grohe C, Gunther A, Koschel D, Kreuter M, Kirsten D, Krogel C, Markart P, Muller-Quernheim J, Neurohr C, Pfeifer M, Prasse A, Schonfeld N, Schreiber J, Wirtz H, Witt C, Costabel U. [Position Paper: Significance of the Forced Vital Capacity in Idiopathic Pulmonary Fibrosis]. *Pneumologie (Stuttgart, Germany)* 2015; 69: 455-458.

72. Prasse A, Probst C, Bargagli E, Zissel G, Toews GB, Flaherty KR, Olschewski M, Rottoli P, Muller-Quernheim J. Serum CC-chemokine ligand 18 concentration predicts outcome in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine* 2009; 179: 717-723.

73. Borie R, Crestani B, Dieude P, Nunes H, Allanore Y, Kannengiesser C, Airo P, Matucci-Cerinic M, Wallaert B, Israel-Biet D, Cadranel J, Cottin V, Gazal S, Peljto AL, Varga J, Schwartz DA, Valeyre D, Grandchamp B. The MUC5B variant is associated with idiopathic pulmonary fibrosis but not with systemic sclerosis interstitial lung disease in the European Caucasian population. *PloS one* 2013; 8: e70621.

74. Vuga LJ, Lindell K, O., Yabes J, Zhang Y, Kaminski N, Gibson K, F., Duncan S, R. Longitudinal Peripheral Blood Protein Profiles Are Predictors of Outcomes In IPF. C14 The golden fingerprint : Molecular characterization of Interstitial lung disease American Thoracic Society; 2014. p. A3925-A3925.

75. Vuga LJ, Milosevic J, Pandit K, Ben-Yehudah A, Chu Y, Richards T, Sciurba J, Myerburg M, Zhang Y, Parwani AV, Gibson KF, Kaminski N. Cartilage oligomeric matrix protein in idiopathic pulmonary fibrosis. *PloS one* 2013; 8: e83120.

76. Dillingh MR, van den Blink B, Moerland M, van Dongen MG, Levi M, Kleinjan A, Wijsenbeek MS, Lupher ML, Jr., Harper DM, Getsy JA, Hoogsteden HC, Burggraaf J. Recombinant human serum amyloid P in healthy volunteers and patients with pulmonary fibrosis. *Pulmonary pharmacology & therapeutics* 2013; 26: 672-676.

77. Homolka J, Ziegenhagen MW, Gaede KI, Entzian P, Zissel G, Muller-Quernheim J. Systemic immune cell activation in a subgroup of patients with idiopathic pulmonary fibrosis. *Respiration; international review of thoracic diseases* 2003; 70: 262-269.

78. Le TT, Karmouty-Quintana H, Melicoff E, Le TT, Weng T, Chen NY, Pedroza M, Zhou Y, Davies J, Philip K, Molina J, Luo F, George AT, Garcia-Morales LJ, Bunge RR, Bruckner BA, Loebe M, Seethamraju H, Agarwal SK, Blackburn MR. Blockade of IL-6 Trans signaling attenuates pulmonary fibrosis. *Journal of immunology (Baltimore, Md : 1950)* 2014; 193: 3755-3768.

79. Oruqaj G, Karnati S, Vijayan V, Kotarkonda LK, Boateng E, Zhang W, Ruppert C, Gunther A, Shi W, Baumgart-Vogt E. Compromised peroxisomes in idiopathic pulmonary fibrosis, a vicious cycle inducing a higher fibrotic response via TGF-beta signaling. *Proceedings of the National Academy of Sciences of the United States of America* 2015; 112: E2048-2057.

80. Krafft E, Laurila HP, Peters IR, Bureau F, Peeters D, Day MJ, Rajamaki MM, Clercx C. Analysis of gene expression in canine idiopathic pulmonary fibrosis. *Veterinary journal (London, England : 1997)* 2013; 198: 479-486.

81. Vasakova M, Striz I, Slavcev A, Jandova S, Kolesar L, Sulc J. Th1/Th2 cytokine gene polymorphisms in patients with idiopathic pulmonary fibrosis. *Tissue antigens* 2006; 67: 229-232.

82. Muhlmann J, Thaler J, Hilbe W, Bechter O, Erdel M, Utermann G, Duba HC. Fluorescence in situ hybridization (FISH) on peripheral blood smears for monitoring Philadelphia chromosome-positive chronic myeloid leukemia (CML) during interferon treatment: a new strategy for remission assessment. *Genes, chromosomes & cancer* 1998; 21: 90-100.

83. Tsoutsou PG, Gourgoulianis KI, Petinaki E, Mpaka M, Efremidou S, Maniatis A, Molyvdas PA. ICAM-1, ICAM-2 and ICAM-3 in the sera of patients with idiopathic pulmonary fibrosis. *Inflammation* 2004; 28: 359-364.

84. Agassandian M, Tedrow JR, Sembrat J, Kass DJ, Zhang Y, Goncharova EA, Kaminski N, Mallampalli RK, Vuga LJ. VCAM-1 is a TGF-beta1 inducible gene upregulated in idiopathic pulmonary fibrosis. *Cellular signalling* 2015; 27: 2467-2473.

85. Kagiyama N, Takayanagi N, Kanauchi T, Ishiguro T, Yanagisawa T, Sugita Y. Antineutrophil cytoplasmic antibody-positive conversion and microscopic polyangiitis development in patients with idiopathic pulmonary fibrosis. *BMJ open respiratory research* 2015; 2: e000058.

86. DePianto DJ, Chandriani S, Abbas AR, Jia G, N'Diaye EN, Caplazi P, Kauder SE, Biswas S, Karnik SK, Ha C, Modrusan Z, Matthay MA, Kukreja J, Collard HR, Egen JG, Wolters PJ, Arron JR. Heterogeneous gene expression signatures correspond to distinct lung pathologies and biomarkers of disease severity in idiopathic pulmonary fibrosis. *Thorax* 2015; 70: 48-56.

87. Craig VJ, Quintero PA, Fyfe SE, Patel AS, Knolle MD, Kobzik L, Owen CA. Profibrotic activities for matrix metalloproteinase-8 during bleomycin-mediated lung injury. *Journal of immunology (Baltimore, Md : 1950)* 2013; 190: 4283-4296.

88. Sokai A, Handa T, Tanizawa K, Oga T, Uno K, Tsuruyama T, Kubo T, Ikezoe K, Nakatsuka Y, Tanimura K, Muro S, Hirai T, Nagai S, Chin K, Mishima M. Matrix metalloproteinase-10: a novel biomarker for idiopathic pulmonary fibrosis. *Respiratory research* 2015; 16: 120.

89. Papaioannou AI, Kostikas K, Manali ED, Papadaki G, Roussou A, Spathis A, Mazioti A, Tomos I, Papanikolaou I, Loukides S, Chainis K, Karakitsos P, Griese M, Papiris S. Serum Levels of Surfactant Proteins in Patients with Combined Pulmonary Fibrosis and Emphysema (CPFE). *PloS one* 2016; 11: e0157789.

90. Foteini Malli PG, Valotassiou Varvara, Irene Tsilioni, Konstantinos Gourgoulianis, Zoe Daniil. Leptin and adiponectin levels in idiopathic pulmonary fibrosis. *European Respiratory Journal* 2011; 38.

91. Wygrecka M, Zakrzewicz D, Taborski B, Didiasova M, Kwapiszewska G, Preissner KT, Markart P. TGF-beta1 induces tissue factor expression in human lung fibroblasts in a PI3K/JNK/Akt-dependent and AP-1-dependent manner. *American journal of respiratory cell and molecular biology* 2012; 47: 614-627.

92. Pardo A, Gibson K, Cisneros J, Richards TJ, Yang Y, Becerril C, Yousem S, Herrera I, Ruiz V, Selman M, Kaminski N. Up-regulation and profibrotic role of osteopontin in human idiopathic pulmonary fibrosis. *PLoS medicine* 2005; 2: e251.

93. Xue J, Kass DJ, Bon J, Vuga L, Tan J, Csizmadia E, Otterbein L, Soejima M, Levesque MC, Gibson KF, Kaminski N, Pilewski JM, Donahoe M, Sciurba FC, Duncan SR. Plasma B lymphocyte stimulator and B cell differentiation in idiopathic pulmonary fibrosis patients. *Journal of immunology (Baltimore, Md : 1950)* 2013; 191: 2089-2095.

94. Collard HR, Calfee CS, Wolters PJ, Song JW, Hong SB, Brady S, Ishizaka A, Jones KD, King TE, Jr., Matthay MA, Kim DS. Plasma biomarker profiles in acute exacerbation of idiopathic pulmonary fibrosis. *American journal of physiology Lung cellular and molecular physiology* 2010; 299: L3-7.

95. Justilien V, Regala RP, Tseng IC, Walsh MP, Batra J, Radisky ES, Murray NR, Fields AP. Matrix metalloproteinase-10 is required for lung cancer stem cell maintenance, tumor initiation and metastatic potential. *PloS one* 2012; 7: e35040.

96. Mariya T, Hirohashi Y, Torigoe T, Tabuchi Y, Asano T, Saijo H, Kuroda T, Yasuda K, Mizuuchi M, Saito T, Sato N. Matrix metalloproteinase-10 regulates stemness of ovarian cancer stem-like cells by activation of canonical Wnt signaling and can be a target of chemotherapy-resistant ovarian cancer. *Oncotarget* 2016.

97. Li J, Yang M, Li P, Su Z, Gao P, Zhang J. Idiopathic pulmonary fibrosis will increase the risk of lung cancer. *Chinese medical journal* 2014; 127: 3142-3149.

98. Ries C. Cytokine functions of TIMP-1. *Cellular and molecular life sciences : CMLS* 2014; 71: 659-672.

99. Song T, Dou C, Jia Y, Tu K, Zheng X. TIMP-1 activated carcinoma-associated fibroblasts inhibit tumor apoptosis by activating SDF1/CXCR4 signaling in hepatocellular carcinoma. *Oncotarget* 2015; 6: 12061-12079.

100. Aaron CP, Schwartz JE, Bielinski SJ, Hoffman EA, Austin JH, Oelsner EC, Donohue KM, Kalhan R, Berardi C, Kaufman JD, Jacobs DR, Jr., Tracy RP, Barr RG. Intercellular adhesion molecule 1 and progression of percent emphysema: the MESA Lung Study. *Respiratory medicine* 2015; 109: 255-264.

101. Reglero-Real N, Alvarez-Varela A, Cernuda-Morollon E, Feito J, Marcos-Ramiro B, Fernandez-Martin L, Gomez-Lechon MJ, Muntane J, Sandoval P, Majano PL, Correas I, Alonso MA, Millan J. Apicobasal polarity controls lymphocyte adhesion to hepatic epithelial cells. *Cell reports* 2014; 8: 1879-1893.

102. Bartolome RA, Sanz-Rodriguez F, Robledo MM, Hidalgo A, Teixido J. Rapid up-regulation of alpha4 integrin-mediated leukocyte adhesion by transforming growth factor-beta1. *Molecular biology of the cell* 2003; 14: 54-66.

103. Kikkawa T, Sato N, Kojika M, Takahashi G, Aoki K, Hoshikawa K, Akitomi S, Shozushima T, Suzuki K, Wakabayashi G, Endo S. Significance of measuring S100A12 and sRAGE in the serum of sepsis patients with postoperative acute lung injury. *Digestive surgery* 2010; 27: 307-312.

104. Qazi BS, Tang K, Qazi A. Recent advances in underlying pathologies provide insight into interleukin-8 expression-mediated inflammation and angiogenesis. *International journal of inflammation* 2011; 2011: 908468.

105. Jones MG, Fabre A, Schneider P, Cinetto F, Sgalla G, Mavrogordato M, Jogai S, Alzetani A, Marshall BG, O'Reilly KM, Warner JA, Lackie PM, Davies DE, Hansell DM, Nicholson AG, Sinclair I, Brown KK, Richeldi L. Three-dimensional characterization of fibroblast foci in idiopathic pulmonary fibrosis. *JCI insight* 2016; 1.

106. Xaubet A, Serrano-Mollar A, Ancochea J. Pirfenidone for the treatment of idiopathic pulmonary fibrosis. *Expert Opin Pharmacother* 2014; 15: 275-281.

107. Raghu G, Selman M. Nintedanib and pirfenidone. New antifibrotic treatments indicated for idiopathic pulmonary fibrosis offer hopes and raises questions. *Am J Respir Crit Care Med* 2015; 191: 252-254.

108. Karimi-Shah BA, Chowdhury BA. Forced vital capacity in idiopathic pulmonary fibrosis--FDA review of pirfenidone and nintedanib. *N Engl J Med* 2015; 372: 1189-1191.

109. Kulkarni T, de Andrade JA, Zhou Y, Luckhardt TR, Thannickal VJ. Alveolar Epithelial Disintegrity in Pulmonary Fibrosis. *American journal of physiology Lung cellular and molecular physiology* 2016: ajplung.00115.02016.

110. Bellaye PS, Kolb M. Why do patients get idiopathic pulmonary fibrosis? Current concepts in the pathogenesis of pulmonary fibrosis. *BMC medicine* 2015; 13: 176.

111. Parker MW, Rossi D, Peterson M, Smith K, Sikstrom K, White ES, Connett JE, Henke CA, Larsson O, Bitterman PB. Fibrotic extracellular matrix activates a profibrotic positive feedback loop. *J Clin Invest* 2014; 124: 1622-1635.

112. Buttner C, Skupin A, Rieber EP. Transcriptional activation of the type I collagen genes COL1A1 and COL1A2 in fibroblasts by interleukin-4: analysis of the functional collagen promoter sequences. *J Cell Physiol* 2004; 198: 248-258.

113. Lin X, Sime PJ, Xu H, Williams MA, LaRussa L, Georas SN, Guo J. Yin yang 1 is a novel regulator of pulmonary fibrosis. *Am J Respir Crit Care Med* 2011; 183: 1689-1697.

114. Ikeda H, Sunazuka T, Suzuki H, Hamasaki Y, Yamazaki S, Omura S, Hatamochi A. EM703, the new derivative of erythromycin, inhibits transcription of type I collagen in normal and scleroderma fibroblasts. *J Dermatol Sci* 2008; 49: 195-205.

115. Lin CH, Yu MC, Tung WH, Chen TT, Yu CC, Weng CM, Tsai YJ, Bai KJ, Hong CY, Chien MH, Chen BC. Connective tissue growth factor induces collagen I expression in human lung fibroblasts through the Rac1/MLK3/JNK/AP-1 pathway. *Biochim Biophys Acta* 2013; 1833: 2823-2833.

116. Rajasekaran S, Vaz M, Reddy SP. Fra-1/AP-1 transcription factor negatively regulates pulmonary fibrosis in vivo. *PLoS One* 2012; 7: e41611.

117. Melboucy-Belkhir S, Pradere P, Tadbiri S, Habib S, Bacrot A, Brayer S, Mari B, Besnard V, Mailleux A, Guenther A, Castier Y, Mal H, Crestani B, Plantier L. Forkhead Box F1 represses cell growth and inhibits COL1 and ARPC2 expression in lung fibroblasts in vitro. *Am J Physiol Lung Cell Mol Physiol* 2014; 307: L838-847.

118. Morgunova E, Yin Y, Jolma A, Dave K, Schmierer B, Popov A, Eremina N, Nilsson L, Taipale J. Structural insights into the DNA-binding specificity of E2F family transcription factors. *Nature communications* 2015; 6: 10050.

119. Logan N, Graham A, Zhao X, Fisher R, Maiti B, Leone G, La Thangue NB. E2F-8: an E2F family member with a similar organization of DNA-binding domains to E2F-7. *Oncogene* 2005; 24: 5000-5004.

120. Maiti B, Li J, de Bruin A, Gordon F, Timmers C, Opavsky R, Patil K, Tuttle J, Cleghorn W, Leone G. Cloning and characterization of mouse E2F8, a novel mammalian E2F family member capable of blocking cellular proliferation. *The Journal of biological chemistry* 2005; 280: 18211-18220.

121. Weijts BG, Bakker WJ, Cornelissen PW, Liang KH, Schaftenaar FH, Westendorp B, de Wolf CA, Paciejewska M, Scheele CL, Kent L, Leone G, Schulte-Merker S, de Bruin A. E2F7 and E2F8 promote angiogenesis through transcriptional activation of VEGFA in cooperation with HIF1. *The EMBO journal* 2012; 31: 3871-3884.

122. Park SA, Platt J, Lee JW, Lopez-Giraldez F, Herbst RS, Koo JS. E2F8 as a Novel Therapeutic Target for Lung Cancer. *Journal of the National Cancer Institute* 2015; 107.

123. Deng Q, Wang Q, Zong WY, Zheng DL, Wen YX, Wang KS, Teng XM, Zhang X, Huang J, Han ZG. E2F8 contributes to human hepatocellular carcinoma via regulating cell proliferation. *Cancer research* 2010; 70: 782-791.

124. Nemoto K, Ikeda A, Yoshida C, Kimura J, Mori J, Fujiwara H, Yokosuka A, Mimaki Y, Ohizumi Y, Degawa M. Characteristics of nobiletin-mediated alteration of gene expression in cultured cell lines. *Biochemical and biophysical research communications* 2013; 431: 530-534.

125. Shimada Y, Kuninaga S, Ariyoshi M, Zhang B, Shiina Y, Takahashi Y, Umemoto N, Nishimura Y, Enari H, Tanaka T. E2F8 promotes hepatic steatosis through FABP3 expression in diet-induced obesity in zebrafish. *Nutrition & metabolism* 2015; 12: 17.

126. Hochstrasser M. Biochemistry. All in the ubiquitin family. *Science* 2000; 289: 563-564.

127. Tanaka Y, Tanaka N, Saeki Y, Tanaka K, Murakami M, Hirano T, Ishii N, Sugamura K. c-Cbl-dependent monoubiquitination and lysosomal degradation of gp130. *Mol Cell Biol* 2008; 28: 4805-4818.

128. Hershko A, Ciechanover A. The ubiquitin system. *Annu Rev Biochem* 1998; 67: 425-479.

129. Cardozo T, Pagano M. The SCF ubiquitin ligase: insights into a molecular machine. *Nat Rev Mol Cell Biol* 2004; 5: 739-751.

130. Deshaies RJ, Joazeiro CA. RING domain E3 ubiquitin ligases. *Annu Rev Biochem* 2009; 78: 399-434.

131. Skowyra D, Craig KL, Tyers M, Elledge SJ, Harper W. F-box proteins function as receptors to recruit phosphorylated substrates to E3 ubiquitin ligase complexes. *Mol Biol Cell* 1997; 8: 2059-2059.

132. Tyers M, Willems AR. One ring to rule a superfamily of E3 ubiquitin ligases. *Science* 1999; 284: 601, 603-604.

133. Skaar JR, Pagan JK, Pagano M. Mechanisms and function of substrate recruitment by F-box proteins. *Nat Rev Mol Cell Biol* 2013; 14: 369-381.

134. Vuga LJ, Ben-Yehudah A, Kovkarova-Naumovski E, Oriss T, Gibson KF, Feghali-Bostwick C, Kaminski N. WNT5A is a regulator of fibroblast proliferation and resistance to apoptosis. *American journal of respiratory cell and molecular biology* 2009; 41: 583-589.

135. Wu W, Dave N, Tseng GC, Richards T, Xing EP, Kaminski N. Comparison of normalization methods for CodeLink Bioarray data. *BMC bioinformatics* 2005; 6: 309.

136. Segal E, Friedman N, Kaminski N, Regev A, Koller D. From signatures to models: understanding cancer using microarrays. *Nature genetics* 2005; 37 Suppl: S38-45.

137. Cenciarelli C, Chiaur DS, Guardavaccaro D, Parks W, Vidal M, Pagano M. Identification of a family of human F-box proteins. *Curr Biol* 1999; 9: 1177-1179.

138. Deng Y, Su Q, Mo J, Fu X, Zhang Y, Lin EH. Celecoxib downregulates CD133 expression through inhibition of the Wnt signaling pathway in colon cancer cells. *Cancer investigation* 2013; 31: 97-102.

139. Parisi F, Ariyan S, Narayan D, Bacchiocchi A, Hoyt K, Cheng E, Xu F, Li P, Halaban R, Kluger Y. Detecting copy number status and uncovering subclonal markers in heterogeneous tumor biopsies. *BMC genomics* 2011; 12: 230.

140. Qi QR, Zhao XY, Zuo RJ, Wang TS, Gu XW, Liu JL, Yang ZM. Involvement of atypical transcription factor E2F8 in the polyploidization during mouse and human decidualization. *Cell cycle (Georgetown, Tex)* 2015; 14: 1842-1858.

141. Ouseph MM, Li J, Chen HZ, Pecot T, Wenzel P, Thompson JC, Comstock G, Chokshi V, Byrne M, Forde B, Chong JL, Huang K, Machiraju R, de Bruin A, Leone G. Atypical E2F repressors and activators coordinate placental development. *Developmental cell* 2012; 22: 849-862.