ORIGINAL ARTICLE

Relationship of DNA Methylation and Gene Expression in Idiopathic Pulmonary Fibrosis

Ivana V. Yang[1,2], Brent S. Pedersen[1], Einat Rabinovich[3], Corinne E. Hennessy[1], Elizabeth J. Davidson[1], Elissa Murphy[1], Brenda Juan Guardela[4], John R. Tedrow[3], Yingze Zhang[3], Mandal K. Singh[3], Mick Correll[5], Marvin I. Schwarz[1], Mark Geraci[1], Frank C. Sciurba[3], John Quackenbush[5], Avrum Spira[6], Naftali Kaminski[4], and David A. Schwartz[1,2]

[1]Department of Medicine, University of Colorado School of Medicine, Aurora, Colorado; [2]Center for Genes, Environment and Health, National Jewish Health, Denver, Colorado; [3]Simmons Center for Interstitial Lung Disease and Department of Medicine, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania; [4]Department of Internal Medicine, Yale School of Medicine, New Haven, Connecticut; [5]Dana-Farber Cancer Institute and Harvard School of Public Health, Boston, Massachusetts; and [6]Department of Medicine, Boston University School of Medicine, Boston, Massachusetts

Abstract

Rationale: Idiopathic pulmonary fibrosis (IPF) is an untreatable and often fatal lung disease that is increasing in prevalence and is caused by complex interactions between genetic and environmental factors. Epigenetic mechanisms control gene expression and are likely to regulate the IPF transcriptome.

Objectives: To identify methylation marks that modify gene expression in IPF lung.

Methods: We assessed DNA methylation (comprehensive high-throughput arrays for relative methylation arrays [CHARM]) and gene expression (Agilent gene expression arrays) in 94 patients with IPF and 67 control subjects, and performed integrative genomic analyses to define methylation–gene expression relationships in IPF lung. We validated methylation changes by a targeted analysis (Epityper), and performed functional validation of one of the genes identified by our analysis.

Measurements and Main Results: We identified 2,130 differentially methylated regions (DMRs; <5% false discovery rate), of which 738 are associated with significant changes in gene expression and enriched for expected inverse relationship between methylation and expression ($P < 2.2 \times 10^{-16}$). We validated 13/15 DMRs by targeted analysis of methylation. Methylation–expression quantitative trait loci (methyl-eQTL) identified methylation marks that control *cis* and *trans* gene expression, with an enrichment for *cis* relationships ($P < 2.2 \times 10^{-16}$). We found five *trans* methyl-eQTLs where a methylation change at a single DMR is associated with transcriptional changes in a substantial number of genes; four of these DMRs are near transcription factors (castor zinc finger 1 [CASZ1], FOXC1, MXD4, and ZDHHC4). We studied the *in vitro* effects of change in CASZ1 expression and validated its role in regulation of target genes in the methyl-eQTL.

Conclusions: These results suggest that DNA methylation may be involved in the pathogenesis of IPF.

Keywords: DNA methylation; gene expression; pulmonary fibrosis; quantitative trait; mapping

Idiopathic pulmonary fibrosis (IPF) appears to result from reprogramming of injured alveolar epithelial cells, which undergo early apoptosis, epithelial–mesenchymal transition, and produce mediators that lead to proliferation of resident fibroblasts and recruitment of fibrocytes. As the extracellular matrix (ECM) expands, myofibroblastic foci develop, resulting in further fibroproliferation in the ECM and the more extensive lung remodeling (1).

(Received in original form August 11, 2014; accepted in final form October 17, 2014)

Supported by National Institutes of Health/National Heart, Lung, and Blood Institute grant R01-HL095393 and Lung Genomics Research Consortium grant RC2-HL101715.

Author Contributions: I.V.Y., M.G., F.C.S., J.Q., A.S., N.K., and D.A.S. designed the study; I.V.Y., B.S.P., E.R., M.C., and J.R.T. analyzed the data; I.V.Y., C.E.H., E.J.D., E.M., B.J.G., Y.Z., and M.K.S. performed laboratory work; M.I.S. and F.C.S. performed clinical and pathological phenotyping of lung specimens; I.V.Y. and D.A.S. wrote the manuscript.

Correspondence and requests for reprints should be addressed to Ivana V. Yang, Ph.D., University of Colorado Denver, 12700 East 19th Avenue, 8611, Aurora, CO 80045. E-mail: ivana.yang@ucdenver.edu

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Am J Respir Crit Care Med Vol 190, Iss 11, pp 1263–1272, Dec 1, 2014

Originally Published in Press as DOI: 10.1164/rccm.201408-1452OC on October 21, 2014
Internet address: www.atsjournals.org

At a Glance Commentary

Scientific Knowledge on the Subject: Idiopathic pulmonary fibrosis (IPF) is currently an untreatable disease of unknown etiology that has a strong genetic component, but is also influenced by exposures to environmental factors, such as cigarette smoke. IPF lung tissue is extensively remodeled and expression of many genes is altered in the diseased lung.

What This Study Adds to the Field: We identified extensive alterations in DNA methylation profiles and strong associations with gene expression changes in diseased lung, both in genes that are known to be important in fibrosis as well as novel candidate genes. These DNA methylation marks represent promising therapeutic targets for IPF.

Gene expression profiling studies have demonstrated that several thousand genes are differentially expressed in the lung parenchyma of individuals with IPF (2, 3). In aggregate, these studies have consistently identified genes associated with ECM formation, degradation, signaling, smooth muscle markers, growth factors, developmental pathways, immunoglobulins, complement, and chemokines (4). The influence of cigarette smoking (5) and the dynamic gene expression changes (4) identified in IPF suggest that epigenetic mechanisms may influence the development of this disease (6). Early studies have demonstrated a link between exposure to tobacco smoke and lung cancer via methylation of CpG islands in cancer genes, such as p16 (7). More recent studies have shown that cigarette smoke has an influence on the epigenome (8–10) and on methylation of specific promoters in genes involved in pathogenesis of IPF, such as *Wnt7a* (11).

Several targeted studies have shown that epigenetic modulation regulates expression of genes involved in pathogenesis of IPF, namely, *cyclo-oxygenase-2* (12), chemokine *IP-10* (13), *Thy-1* (*CD90*) (14), *p14* (*ARF*) (15), and *α-smooth muscle actin* (*α-SMA*) (16, 17). Two genomic studies of DNA methylation in IPF lung have revealed extensive DNA methylation changes within CpG islands (18, 19). These early studies suggested that DNA methylation may influence the expression of specific genes in the IPF lung; however, both studies used small sample sizes, only partially characterized the relationship of DNA methylation and gene expression, and were unable to assess methylation of less dense CpG areas near CpG islands (CpG shores) that have recently been shown to be associated with tissue-specific gene expression, cancer, hematopoietic cell lineage commitment, and stem cell pluripotency (20–22). In our study, we used comprehensive high-throughput arrays for relative methylation (CHARM) methodology (23) to interrogate 4.6 million CpG sites distributed across the human genome with gene expression changes in lung tissue of 94 subjects with IPF and 67 control subjects to more comprehensively study the association between DNA methylation and gene expression in IPF. Partial results of these studies have been reported in the form of an abstract (24).

Methods

Subjects and Tissue Samples

All human tissue was collected after appropriate ethical review for the protection of human subjects through the National Heart, Lung, and Blood Institute–sponsored Lung Tissue Research Consortium. All subjects met American Thoracic Society/European Respiratory Society criteria for diagnosis of IPF (25). Please refer to the online supplement for population description. Lung tissue specimens from lower ($n = 52$), upper ($n = 34$), and middle/lingula ($n = 8$) lobes from subjects with IPF and with normal histology from lower ($n = 26$), upper ($n = 34$), and middle/lingula ($n = 7$) lobes were obtained from the Lung Tissue Research Consortium.

Genomic DNA Methylation and Gene Expression Profiling

DNA and RNA were isolated using standard methods. We followed CHARM protocols developed by others using Nimblegen CHARM array design (23, 26) with additional quality control metrics. Gene expression profiling was performed using Agilent's human gene expression microarrays (GE 4 × 44 k v2 or G3 Sure print 8 × 60 k formats; Agilent Technologies, Santa Clara, CA).

Validation of DNA Methylation Data

Methylation of individual CpG motifs within selected differentially methylated regions (DMRs) identified using CHARM was examined in a subset of subjects (52 IPF and 48 control) using MassARRAY EpiTYPER assays and primers designed in the EpiDesigner software (Sequenom, San Diego, CA; *see* Table E1 in the online supplement). For analysis of castor zinc finger (CASZ) 1 methylation in alveolar type II cells (27), we used pyrosequencing on the PyroMark Q96 MD sequencer (Qiagen, Germantown, MD).

DNA Methylation Data Analysis

Data quality for CHARM arrays was assessed using the quality control functions in the charm R package (28), principal components analysis, and probabilistic estimation of expression residuals factors (29, 30). Normalized percent methylation estimates (range, 0–1) at each of the 2.1 million probes were used in all downstream analyses after logit transformation. The R package limma (31) was used to fit linear models for methylation data and *P* values were based on the moderated *t* statistic. From this, DMRs were identified from those *P* values using comb-p (32) with a window size of 300 bases and a corrected *P* value of less than 0.05 (corresponding to $\alpha < 5\%$ after correcting for the number of possible DMRs). This resulted in regions of adjacent probes with low *P* values that stand up to genome-wide correction. Regions were annotated to nearest gene and CpG island from human genome version hg18 with CruzDB (33).

Gene Expression Data Analysis

Expression array data were normalized using cyclic loess (34). The R package, limma (31), was used to fit linear models for expression data and *P* values were based on the moderated *t* statistic followed by false discovery rate (FDR) correction to obtain q values.

Quantitative Trait Loci Analysis

Quantitative trait loci (QTL) models were run in Matrix expression QTL (eQTL) (35). Analysis was performed on 1,315 differentially expressed genes in subjects with IPF compared with control subjects (5% FDR and more than twofold change) and 50% of CHARM probes with the largest variance across the entire dataset (94,908 local or *cis* [<1 megabase (Mb)] and 1,183,908,737 distant or *trans* [>1 Mb]

loci). We fit expression data to methylation data and adjusted for disease status, age, sex, smoking, and CHARM probabilistic estimation of expression residual factors. Significant relationships of methylation marks and gene expression were identified after FDR correction.

CASZ1 siRNA Treatment of Calu-3 Cells

Calu-3 human airway epithelial cells were grown and treated with short interfering RNA (siRNA) using standard protocols, as outlined in the online supplement. Genomic expression of four biological replicates of CASZ1 siRNA and off-target siRNA was determined using Agilent 8 × 60 k S3 SurePrint arrays (Agilent Technologies).

CASZ1 Immunohistochemistry

Standard immunohistochemical staining protocols were followed, as described in the online supplement.

Lung Genomics Research Consortium Data

Genomic data (CHARM methylation arrays and Agilent gene expression arrays), as well as associated clinical data, are available for download on the Lung Genomics Research Consortium Web site (https://www.lung-genomics.org/research/).

Results

Demographic Characteristics

Table 1 summarizes demographic and clinical characteristics of the subjects with IPF and nondiseased control subjects. The IPF cohort is composed of more males than the control subjects. Patients with IPF on average have smoked fewer cigarettes than control subjects, but there is substantial variability in pack-years in both groups. Not surprisingly, diseased individuals were significantly more likely to be treated with corticosteroids alone or in combination with immunosuppressive agents. As expected, the IPF cohort is clinically characterized by decreased lung function, increased HRCT evidence of honeycombing (but not emphysema), and significant microscopic honeycombing and fibroblastic foci on pathology.

Widespread DNA Methylation Changes Are Present in IPF Lung

Comparing 94 subjects with IPF to 67 control subjects, we identified 2,130 genome-wide significant DMRs ($\alpha < 5\%$ after correcting for the number of possible DMRs) after adjusting for age, sex, pack-years of smoking, and technical variables (both known and inferred; Figure 1 and Figure E1 and Table E2 in the online supplement). For 90% of the regions, directionality of methylation change is consistent across all probes that constitute the DMR (Table E2).

The median length of DMRs is 332 base pairs (range, 185–2046 base pairs). Of the 2,130 DMRs, 923 (43%) are hypomethylated in subjects with IPF compared with control subjects. Median absolute change in methylation is 5.7% between subjects with IPF and control subjects at each DMR (range, 0.1–17%). The majority of the DMRs (n = 1,522 [71%]) are located within gene bodies (intron, exon, and 3′ untranslated region [UTR] and 5′ UTR), whereas only a small number (n = 203 [10%]) are located in the gene promoter (within 5 kb of the transcription start site [TSS]; Figure 1B). The 2,130 DMRs are associated with 1,514 unique genes; most genes have one DMR, but some have multiple DMRs, such as the

Table 1. Subject Demographics and Clinical Characteristics of the Lung Tissue Research Consortium Cohort Used in the Study to Identify Genomic Methylation and Expression Changes

	Controls	IPF	*P* Value
n	67	94	
Age, mean (SD)	64.0 (10.1)	64.8 (8.4)	0.57*
Sex, % male	49.3	67.0	0.034†
Race, % white	97.0	94.7	0.70*
Smoker, n (%)			0.60*
Current	2 (3)	1 (1)	
Former	43 (64)	65 (69)	
Never	22 (33)	28 (30)	
Pack-years, mean (SD)‡	37.2(35.7)	25.9(19.9)	0.034*
Medication use within 30 d of biopsy, n (%)			
None	60 (90)	23 (25)	8.2×10^{-11}
Corticosteroids	2 (3)	19 (20)	
Immunosuppressants	0 (0)	0 (0)	
Corticosteroids and immunosuppressants	0 (0)	49 (52)	
Unknown	5 (7)	3 (3)	
Pre-BD FVC, % predicted (SD)	96.9 (11.7)	71.9 (16.2)	1.1×10^{-20}*
D_{LCO}, % predicted (SD)	85.1 (17.6)	45.4 (16.6)	3.0×10^{-29}*
Lobe-specific honeycombing (HRCT scan), n (%)§			0.023†
0 = Normal/none	49 (73)	60 (64)	
1 = Mild (1–25%)	3 (4)	20 (21)	
2 = Moderate (26–50%)	2 (3)	6 (6)	
3 = Marked (51–75%)	0 (0)	2 (2)	
4 = Severe (>75%)	0 (0)	2 (2)	
Unknown	13 (19)	4 (4)	
Honeycombing score (HRCT scan)‖ (SD)	0.13 (0.44)	2.94 (4.21)	2.6×10^{-6}*
Percent emphysema (HRCT scan) (SD)	0.71 (1.07)	0.99 (1.82)	0.36*
Lobe–specific lung tissue pathology, microscopic honeycombing, n (%)¶			8.7×10^{-10}*
0 = Unscored/not present	45 (100)	31 (49)	
1 = Rare	0 (0)	5 (9)	
2 = Present	0 (0)	27 (43)	
Lobe-specific lung tissue pathology, fibroblastic foci score, n (%)¶			2.7×10^{-10}*
0 = Unscored/not present	45 (100)	29 (47)	
1 = Rare	0 (0)	21 (33)	
2 = Present	0 (0)	13 (20)	

Definition of abbreviations: BD = bronchodilator; D_{LCO} = diffusing capacity of carbon monoxide; HRCT = high-resolution computed tomography; IPF = idiopathic pulmonary fibrosis.
*Assessed by two-tailed *t* test.
†Assessed by Fisher's exact test.
‡Average for current and former smokers.
§Extent of honeycombing in the lobe that DNA and RNA were extracted from, as determined by HRCT scan. All patients with IPF had evidence of honeycombing in at least one lung lobe.
‖Sum of honeycomb scores for all lobes.
¶Data available on 108 subjects.

CASZ1 gene with 24 intronic DMRs. As observed in colon cancer (20), the majority of the IPF DMRs (n = 1,285 [60%]) are in CpG island shores (1–3,000 bases from the island), whereas only a small number of DMRs (n = 98 [5%]) are within the islands themselves (Figure 1C).

To internally validate our findings, EpiTYPER assays were done on 15 DMRs, selected based on larger methylation changes (range, 7–12%) and to include both genes that have been implicated in pathogenesis of IPF (*NOTCH1*, *FBXO32*, and *TOLLIP*) and those with no previous association with fibroproliferative lung disease (*CASZ1*, *SOX7*, and *TRIM71*). Average methylation across all CpGs in the DMR is significantly different and in the same direction as determined by CHARM analysis in IPF compared with controls in 13 out of the 15 DMRs (Figure 1D and Table E3). Of the two DMRs that did not validate, *NPAS2* has methylation changes in the same direction as CHARM (but the difference did not reach statistical significance), whereas *FBXO32* has no measurable change in methylation between cases with IPF and control subjects using MassARRAY.

We also intersected the 2,130 IPF-associated DMRs with recently defined IPF genome-wide association study (GWAS) loci (36, 37) and identified methylation changes in genes within 5 of the 10 IPF GWAS loci: 4q22 (HERC3; one DMR in the promoter [937 bases upstream of the TSS]); 6p24 (*DSP*; two intronic DMRs); 11p15 (*TOLLIP*; two intronic DMRs and *MOB2*/*IFITM10*; three intronic DMRs); 13q34 (*ATP11A*; 13 intronic and 1 exon/3′ UTR DMRs and MCF2L; six intronic DMRs); and 19p13 (*DPP9*; two intronic DMRs).

Finally, we performed an exploratory analysis to examine whether any of the 2,130 IPF-associated DMRs are influenced by medication use within the IPF group. No DMRs were influenced by corticosteroid use alone (multiple comparison–adjusted $\alpha <$ 0.05), and five of them (*ESRRG*, *FNBP1*, *JAG1*, *SEPT9*, and *SPTBN2*) were influenced by combined use of corticosteroids and immunosuppressants. However, this analysis is limited by the number of samples in each group, and we did not have the power to arrive at any definitive conclusions.

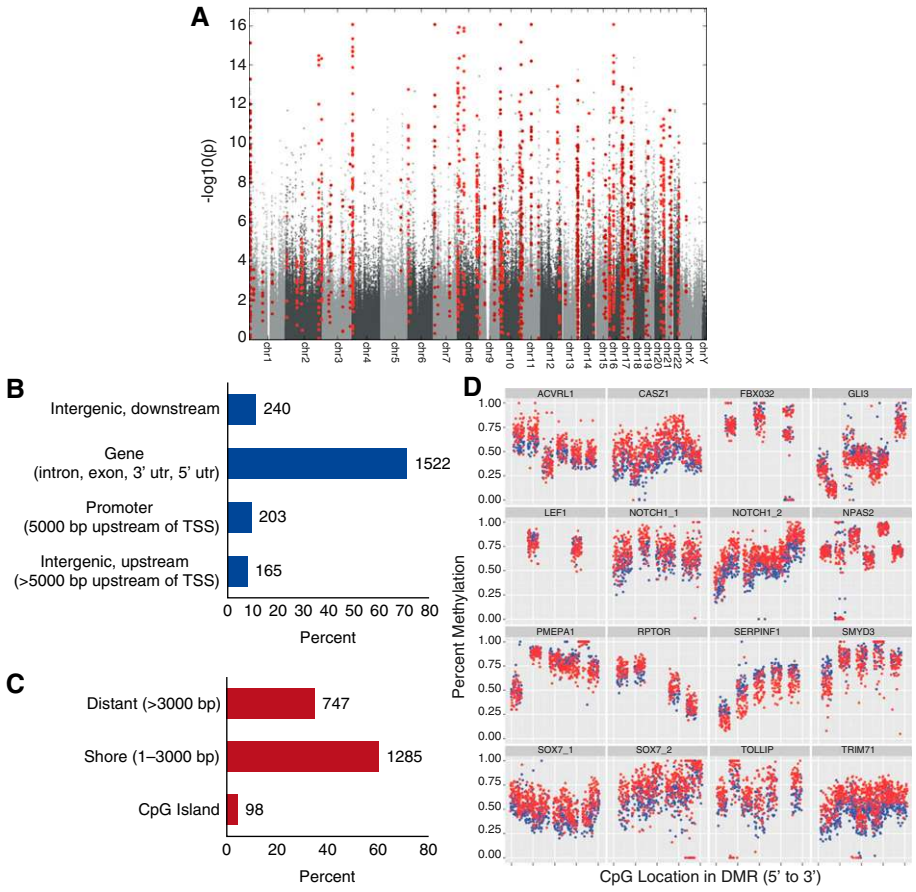

Figure 1. Differentially methylated regions (DMRs) are associated with idiopathic pulmonary fibrosis (IPF) after controlling for age, sex, pack-years of smoking, technical variables, and batch effects. (*A*) Manhattan plot of the adjusted *P* values for disease status (IPF/control) from the linear model. DMRs are identified as significant ($\alpha < 5\%$ after correcting for the number of possible DMRs) using the linear model: % methylation ~ disease + sex + age + pack-years + technician + signal_strength + [five probabilistic estimation of expression residual factors]. Each *dot* represents a *P* value for a probe on the comprehensive high-throughput arrays for relative methylation (CHARM) array that has been adjusted by the significance of neighboring probes within 300 bases according to their correlation. *Highlighted in red* are probes within 127 significant DMRs with most pronounced methylation changes (absolute change ⩾ 10%) between cases and control subjects. Heatmap of all 2,130 DMRs is shown in Figure E1. Genomic distribution of 2,130 DMRs by relationship to (*B*) gene and (*C*) CpG island. (*D*) Internal validation of 15 selected DMRs (two PCR products are included for the DMR in SOX7) on the Epityper platform. Plotted on the *y axis* are percent methylation values versus CpG site on the *x axis*. *Red dots* represent IPF samples; *blue dots* are controls.

Relationship between DNA Methylation and Gene Expression in IPF Lung

Consistent with previous investigations of gene expression in IPF lung, a large number—13,251 unique genes—were significantly differentially expressed genes (5% FDR). Of these genes, 738 had a DMR within 5 kb of the gene (Figure 2; Table E4). As expected, hypomethylated DMRs were associated with increased gene expression, and hypermethylated DMRs were associated with decreased gene expression (69%; $P < 2.2 \times 10^{-16}$; *blue dots* in Figure 2); however, a number of methylation changes were not associated with expected gene expression differences (*red dots*).

To functionally characterize 738 genes with expression and methylation changes, we examined canonical pathways, upstream regulators, and protein–protein networks in the Ingenuity Pathway Analysis. Among the most enriched canonical pathways were several that have been implicated in the pathogenesis of IPF, namely, CXCR4 signaling (Benjamini-Hochberg [B-H] multiple comparison–adjusted $P = 7.8 \times 10^{-3}$) (38), thrombin signaling (B-H $P = 9.1 \times 10^{-3}$) (39), Wnt/β-catenin signaling (B-H $P = 0.017$) (40), vascular endothelial growth factor signaling (B-H $P = 0.029$) (41), and epithelial adherens junction signaling (B-H $P = 0.044$) (42) (Table E5).

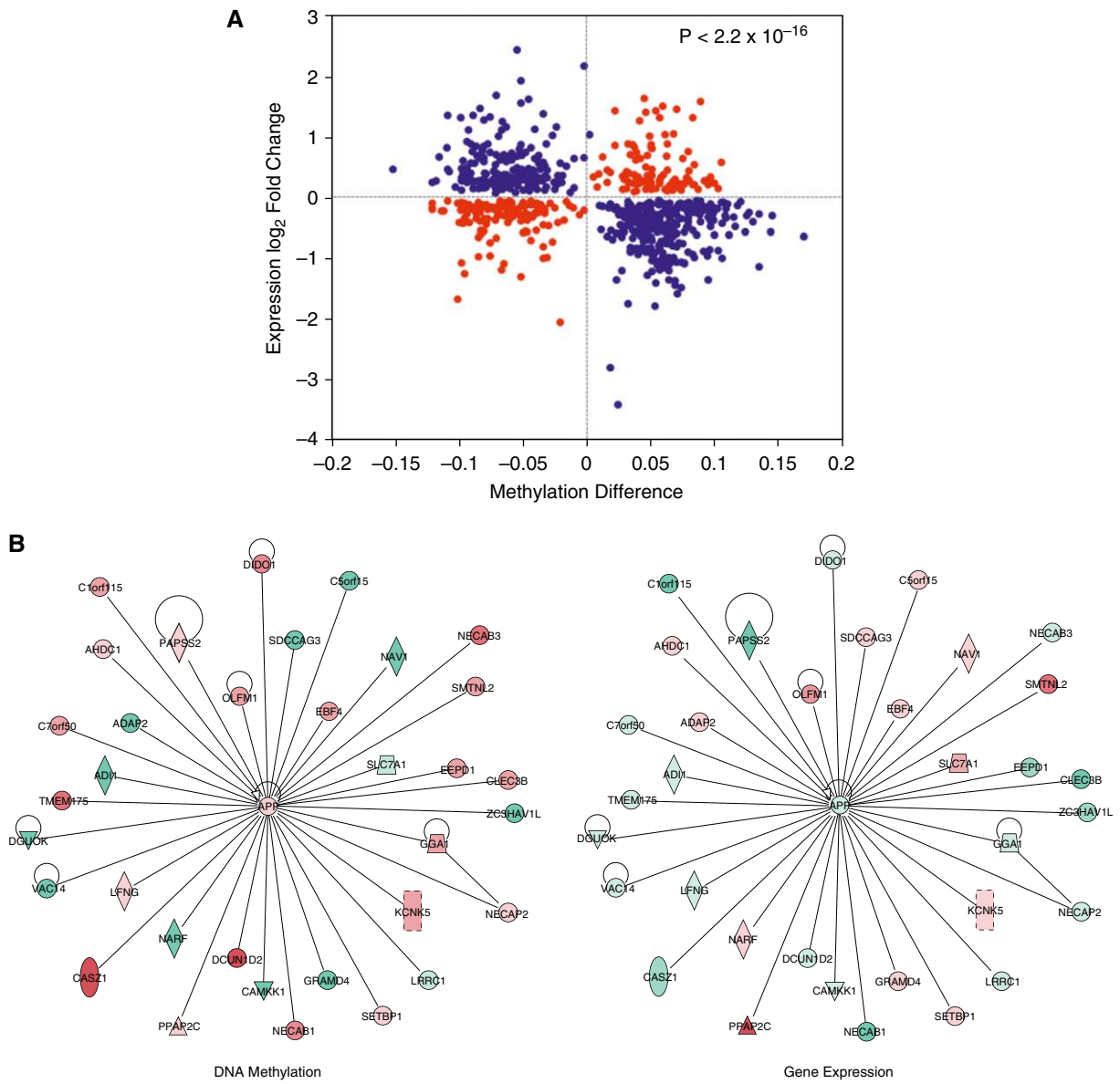

Figure 2. Expression changes in genes within 5 kilobases of the nearest differentially methylated region (DMR) associated with idiopathic pulmonary fibrosis (IPF). (*A*) Plotted are expression data for 738 genes that have a DMR within 5 kilobases of the gene (defined from the transcription start site to the last base of the last exon). The *x-axis* methylation difference is represented by the mean percent methylation difference in IPF compared with control; *y axis* expression difference is represented by the mean fold change in IPF compared with control (on the log base 2 scale). The *blue dots* represent hypomethylated genes that were associated with increased gene expression and hypermethylated genes were associated with decreased gene expression. The *red dots* represent methylation changes that were not associated with expected gene expression differences. The *P* value is the binomial *P* value for enrichment of expected methylation–gene expression relationships (*blue dots*) in the dataset. (*B*) The most significant protein–protein network (score = 54) identified by Ingenuity Pathway Analysis (IPA) demonstrates that a number of genes with differential methylation and expression in IPF lung are known to have direct interactions (*solid lines*) with the amyloid β precursor using data in the Ingenuity Knowledge Base. *Green* indicates lower methylation or expression in IPF; *red* indicates higher methylation or expression in IPF compared with control lung. Methylation values are colored based on percent methylation change between IPF and controls; colors of expression values are based on $\log_2$(fold change) between IPF and controls. Molecule shapes: *oval* = transcriptional regulator; *diamond* = enzyme; *dashed line rectangle* = channel; *up triangle* = phosphatase; *down triangle* = kinase; *trapezoid* = transporter; *circle* = other. This analysis was restricted to only direct relationships. The network score is based on the hypergeometric distribution, and is calculated with the right-tailed Fisher's exact test to identify enrichment of differentially methylated/expressed genes in the network relative to IPA database. The other nine networks with scores of greater than 30 are shown in Figure E2.

Analysis of binding motifs in promoters of the 738 genes identified overrepresentation of regulators of lung development, specifically, β-catenin ($P = 2.4 \times 10^{-5}$), GLI1 ($P = 8.6 \times 10^{-5}$), and FOXC2 ($P = 6.3 \times 10^{-6}$); FOXC2 is a regulator of mesenchymal tissue development that induces expression of integrins (43), which are key mediators of fibroproliferation (44) (Table E6). Network analysis identified 10 high-scoring (score > 30) networks in the expression–methylation dataset (Figure 2B

and Figure E2). The most significant network centers around amyloid β precursor (APP) protein and contains the transcription factor, CASZ1 (Figure 2B). APP, or presenilin 2, is best known for its role in Alzheimer's disease, and is one of the substrates for γ secretase; other substrates include Notch, E-cadherin, and CD44 (45). However, APP is known to be expressed in the lung, is alternatively spliced in lung cancer (46), and has been shown to bind directly to transforming growth factor-β1 (47). The rest of the networks contain genes related to development (FZD8, GLI3, JAG2, WNT5B, WNT10A), cilium (CELSR1, FOXJ1, RFX2, RPGRIPL1), tight junctions (CTNND2, SLDN5, SLDN11, TJP2), as well as those with IPF-associated genetic variants (DSP, TOLLIP) (Figure E2). Taken together, this analysis identified a number of biologically relevant methylation–expression relationships in IPF lung.

Methylation-eQTL Analysis

To further explore the relationship between methylation and gene expression, we performed a methylation-eQTL (methyl-eQTL) analysis, evaluating the association between epigenetic loci and individual differences in quantitative levels of differentially expressed genes in subjects with IPF compared with control subjects (Figure 3). The goal of this analysis was to identify methylation changes that affect most prominently differentially expressed genes in IPF lung; it was performed on 1,315 most differentially expressed genes in subjects with IPF compared with control subjects (5% FDR and more than twofold change) and 50% of CHARM probes with the highest variance across the entire dataset. At FDR less than 0.1%, 172/1,315 genes have significant relationships of methylation and expression; on average, each gene has eight methylation marks associated with its expression (Table E7). We observed mostly changes in expression that are opposite in direction with DNA methylation (*blue dots*), and detected only a small number of statistically significant changes with the same directionality (*red dots*). Although we observed an enrichment in *cis* relationships (within 1 Mb of the gene; χ^2 $P < 2.2 \times 10^{-16}$), we also identified methylation marks that regulate large groups of genes (*trans* regulation; master regulators of expression). Four out of the five most significant *trans*-regulating methylation marks are near transcription factors (CASZ1, FOXC1, MXD4, ZDHHC4). Analysis of promoters of genes that are putatively regulated by methylation changes in CASZ1 using the ENCODE CHIP-seq data (48) showed enrichment in the EZH2 ($P < 1 \times 10^{-7}$), a methyltransferase and component of the polycomb repressor complex 2 that plays an essential role in the epigenetic maintenance of the H3K27me3 repressive chromatin mark. EZH2 is known to mediate epigenetic silencing of CASZ1 in neuroblastomas (49). Our results are consistent with this relationship; EZH2 mRNA is up-regulated by 1.8-fold ($q = 9 \times 10^{-11}$), whereas CASZ1 is down-regulated by 1.7-fold ($q = 3.3 \times 10^{-10}$) in IPF lung.

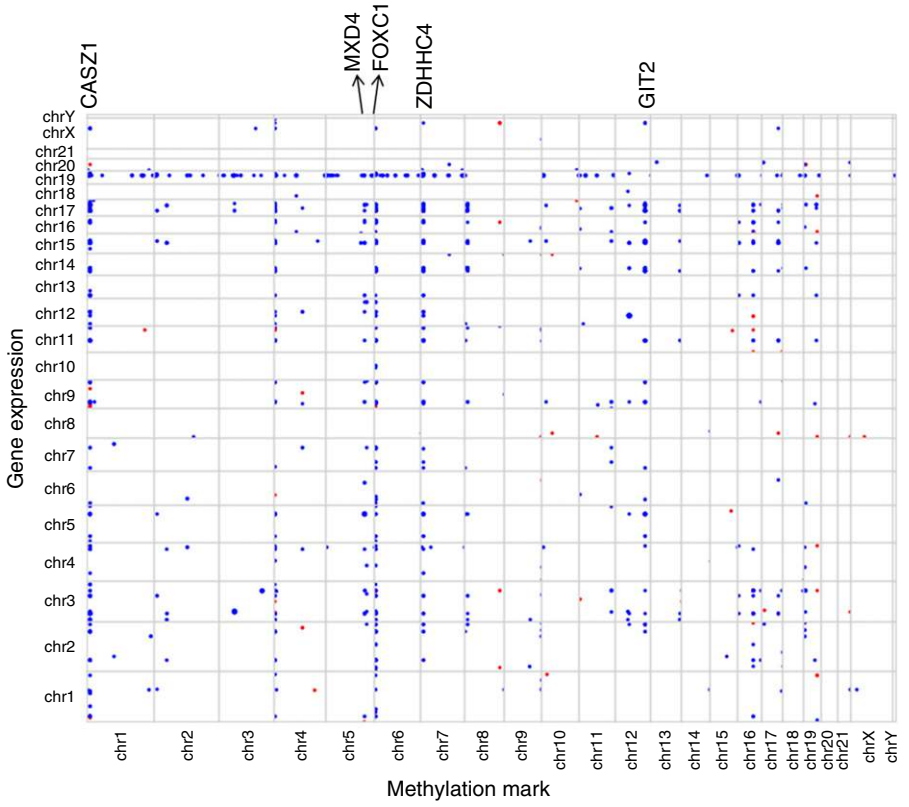

Figure 3. Methylation–expression quantitative trait loci (methyl-eQTL) mapping evaluates the association between epigenetic loci (*x axis*) and individual differences in quantitative levels of expressed genes (*y axis*) in idiopathic pulmonary fibrosis (IPF) cases (n = 94) versus control subjects (n = 67). Statistically significant expression changes that are associated with methylation differences in IPF compared with controls. Analysis was performed on 1,315 differentially expressed genes in IPF compared with controls (5% false discovery rate [FDR] and more than twofold change) and 50% of CHARM probes with most variable probes across the entire dataset. The model used was expression ~ methylation + disease + sex + age + pack-years. On this plot of significant expression changes (*y axis*) that are associated with methylation differences (*x axis*), each *dot* represents a significant difference in expression being affected by methylation (<0.1% FDR) after adjusting for disease, sex, age, and pack-years of smoking. Methylation changes that control expression of genes within 1 megabase (Mb) of the gene are referred to as *cis*-QTLs, whereas methylation changes that regulate expression of more distant (>1 Mb) genes are *trans*-QTLs. *Blue dots* represent inverse relationships of methylation and expression (negative T sum) and *red dots* are positive relationships of methylation and expression (positive T sum). The size of the *dot* is proportional to the *P* value of the significance of the methylation–expression relationship. *Vertical lines* of expression changes indicate methylation marks that regulate large groups of genes (master regulators of expression). The five most significant *trans* lines are annotated with the name of the gene nearest the differentially methylated region (DMR). Modules of genes the expression of which is regulated by methylation marks in the five *trans* methyl-eQTLs highlighted in (*A*) are shown in Figure E3.

The Role of the Transcription Factor, CASZ1, in Regulating Gene Expression in Epithelial Cells

Our methyl-eQTL analysis identified several methylation marks near CASZ1 that are predicted to regulate expression of 63

unique genes in *trans* (Table E7). In the absence of tools to selectively alter methylation status of methylation marks near the CASZ1 gene, we modified expression of CASZ1 by siRNA and studied the effect of reduced CASZ1 expression on expression levels of the 63 *trans* genes in Calu-3 human airway epithelial cell line. Treatment with a pool of four siRNA duplexes for CASZ1 compared with scrambled control pool resulted in an average of 3.9 ($\pm$0.2)-fold down-regulation of CASZ1 in four biological replicates. A total of 21/63 genes predicted to be regulated by CASZ1 methylation are differentially expressed at FDR less than 10% (Table E8A); comparing this to genomic results (6,003 differentially expressed out of 50,559 of all probes on the array; Table E8B) demonstrates a significant enrichment (χ^2 $P = 1.9 \times 10^{-5}$) in predicted CASZ1 targets. These data indicate that change of expression of CASZ1 by methylation is a plausible mechanism in controlling expression of the genes predicted as targets by methyl-eQTL analysis.

Localization of Expression and Methylation of the CASZ1 Transcription Factor in IPF Lung

Given the unknown role of CASZ1 in pathogenesis of pulmonary fibrosis, we pursued further studies with this gene. We first performed immunohistochemistry to localize expression of this transcription factor in normal and IPF lung. In histologically normal lung tissue, CASZ1 is expressed in tracheal, bronchial, and airway epithelia with strong nuclear localization (Figures 4A and 4B), as well as in alveolar type II cells and alveolar macrophages (Figure 4C). In IPF lung, there is a marked overall decrease in airway epithelial expression of CASZ1, with a substantial amount of heterogeneity in expression; some airways in fibrotic areas express reduced levels of this transcription factor (Figure 4D), whereas others do not express it at all (Figure 4E). On the other hand, alveolar type II cells in fibrotic areas (alveolar cysts) (50) appear to express higher concentrations of CASZ1 (Figure 4F).

To address the issue of cell specificity of DNA methylation, especially given differences in expression pattern of CASZ1 in airway versus alveolar epithelia in IPF lung, we measured methylation status of the same CASZ1 DMR validated by Epityper in IPF lung (Chr1: 10,647,117–10,647,381) using pyrosequencing in alveolar type II cells from seven subjects with IPF and four control subjects. We demonstrated an average 3.5% decrease in methylation across eight CpG sites within this DMR in IPF lung compared with controls ($P = 2.6 \times 10^{-3}$; Table E9), consistent with increased expression of CASZ1 in alveolar epithelial cells (Figure 4). These results underscore the importance of examining cell-specific methylation and gene expression.

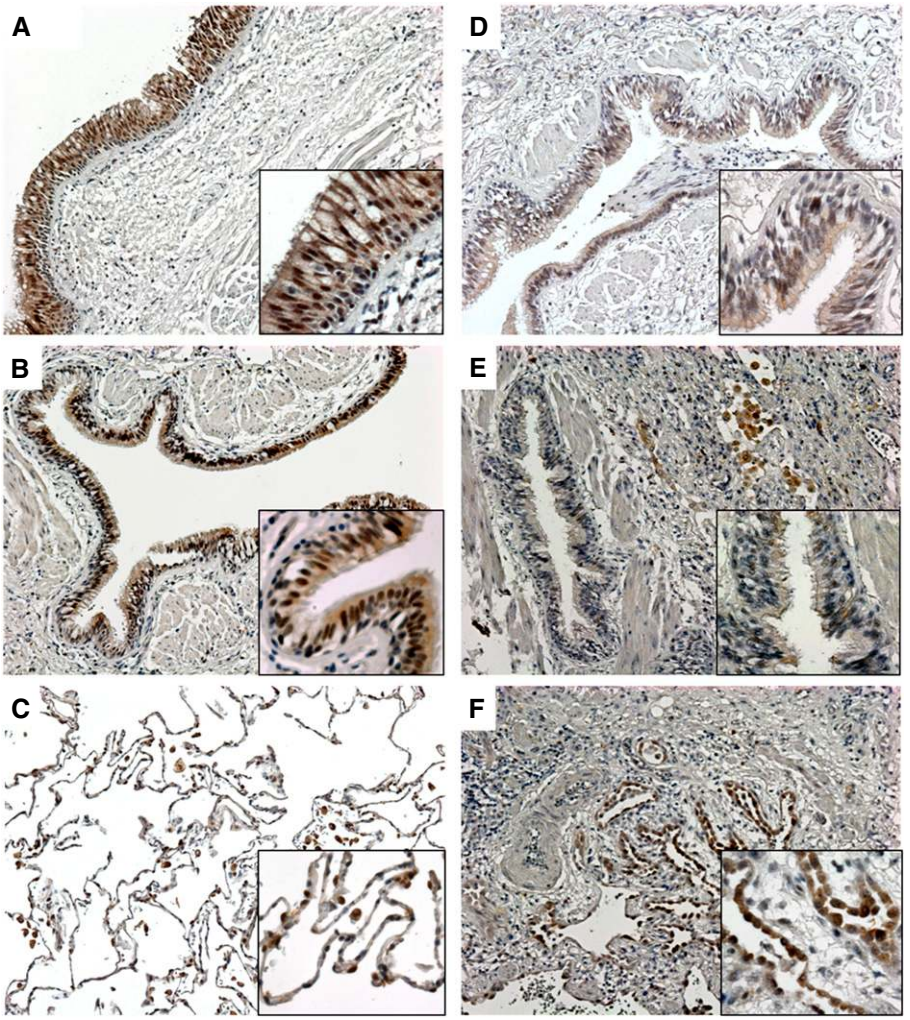

Figure 4. Localization of castor zinc finger (CASZ) 1 expression in idiopathic pulmonary fibrosis (IPF) lung. Immunohistochemical staining of normal (*A–C*) and IPF (*D–F*) lung tissue reveals a decrease in expression of CASZ1 in airway epithelial cells in IPF compared with control lung accompanied by an increase in expression in alveolar type II cells in alveolar cysts in IPF lung. Tissue sections were counterstained with hematoxylin. Images were taken at 10× magnification (40× for *insets*).

Discussion

Our genomic analysis of the IPF lung revealed a large number of disease-associated methylation changes that also affect gene expression. Methylation marks are predominantly located in gene bodies and CpG island shores, and many of the methylation–expression changes follow the expected relationship. Methyl-eQTL analysis further focused our attention on biologically relevant methylation marks, including relevant canonical pathways, protein–protein networks, and transcription factors.

Epigenetics marks can be influenced both by the environment (51) and genetics (52, 53), and it is likely that both factors contribute to widespread epigenetic changes observed in IPF lung. An area that needs further exploration is the influence of environmental exposures, such as cigarette

smoke on IPF-associated epigenetic marks identified in the current study. Although other studies have examined the effect of cigarette smoke on methylation profiles in the airway epithelium (10), no study has done this comprehensively for the alveolar epithelium. Our results demonstrating that half of the IPF GWAS loci have changes in methylation also support the possibility that genetic risk factors may influence DNA methylation in IPF lung. However, the resolution of the CHARM array prevented us from examining the interaction of specific genetic variants and methylation marks in the IPF GWAS loci. Further analysis will be required to definitively address the question of how a combination of environmental exposures, genetic variants, and methylation marks contributes to altered expression and disease phenotypes.

Given that the majority of the methylation changes we identified are outside of CpG islands, and that published studies of DNA methylation in IPF (18, 19) examined only CpG islands using different techniques for assessment of methylation marks (methylated DNA immunoprecipitation [18] and bisulfite conversion [19]), we expected limited gene-level overlap (~7% of genes identified in our study are also in the studies by Rabinovich and colleagues [18] or Sanders and colleagues [19]). However, differential methylation of several genes previously associated with disease (four collagens, *HDAC4*, *NOTCH1*, *PDGF*, *SERPINF1*, and *TOLLIP*), as well as the novel candidate gene, *CASZ1*, was observed in all studies.

Approximately half of the differentially methylated genes are also differentially expressed. These results are in contrast to the previous observation of Sanders and colleagues (19) of the very limited overlap in methylation and expression changes in IPF. One of the possible explanations for this discrepancy is the fact that methylation changes in CpG island shores have recently been shown to be more regulatory than those within islands (20). However, only 10–15% of genes that are more than twofold up- or down-regulated in IPF lung have significant associations with methylation in our analysis. This is consistent with the fact that mechanisms other than DNA methylation regulate gene expression (54).

Using the simple overlap analysis, we observed a statistically significant enrichment for gene expression–methylation relationships opposite in directionality, but also identified a number of genes with the same direction of methylation and expression changes, consistent with findings of others (55). In contrast, our analysis of methylation–expression relationships using QTL mapping identified, almost exclusively, relationships with opposite directionality of methylation and expression changes. Four out of the five most significant *trans*-regulating sets of methylation marks are near transcription factors, suggesting that epigenetic regulation of gene expression in IPF is at least partially mediated by changes in expression of transcription factors that, in turn, regulate expression of downstream genes.

CASZ1 is a transcription factor without a previously established role in the pathogenesis of IPF. CASZ1 plays a role in the development of cardiac and skeletal muscle (56) and blood vessel formation (57). It promotes vascular assembly and morphogenesis by binding to an intronic element in EGFL7 (57), a secreted angiogenic factor that binds to ECM, and has a putative role in Notch signaling (58). Given the prominent role of ECM deposition and recapitulation of developmental pathways in IPF lung, CASZ1 is a strong candidate for further studies in the context of fibroproliferation.

Moreover, identification of the histone methyltransferase, EZH2, as a putative mediator of methylation changes in CASZ1 on regulation of gene expression in IPF, consistent with its known role (49), provides further support for additional studies on the role of CASZ1 in fibroproliferation, with a potential therapeutic option for alteration of gene expression in IPF, given further studies. In fact, HDAC inhibitors have been shown to decrease binding of H3K27me3 and polycomb repressor complex 2 to *CASZ1* gene (49). The main limitation of currently available epigenetic-based therapies is lack of specificity for gene(s) of interest; however, locus-specific therapies are currently being developed (59, 60).

In summary, our analysis identified methylation–gene expression relationships within genes that are either involved in fibroproliferation or are feasible candidates in this process. The main limitation of our study is the analysis of a mixture of cells in lung tissue; further studies in specific cell types, such as airway and alveolar epithelial cells or fibroblasts, are needed. Some of the small methylation changes we identified in the mixture of cells are likely reflective of larger methylation changes in pure cell populations. Another limitation is the use of normal tissue adjacent to lung cancer as control samples; methylation patterns in these tissues may be abnormal due to field cancerization (61). Despite these limitations, identification of extensive epigenetic changes and associated gene expression changes in IPF lung suggests that epigenetic therapies should be explored for this devastating disease. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

References

1. King TE Jr, Pardo A, Selman M. Idiopathic pulmonary fibrosis. *Lancet* 2011;378:1949–1961.
2. Yang IV, Coldren CD, Leach SM, Seibold MA, Murphy E, Lin J, Rosen R, Neidermyer AJ, McKean DF, Groshong SD, *et al*. Expression of cilium-associated genes defines novel molecular subtypes of idiopathic pulmonary fibrosis. *Thorax* 2013;68:1114–1121.
3. Zuo F, Kaminski N, Eugui E, Allard J, Yakhini Z, Ben-Dor A, Lollini L, Morris D, Kim Y, DeLustro B, *et al*. Gene expression analysis reveals matrilysin as a key regulator of pulmonary fibrosis in mice and humans. *Proc Natl Acad Sci USA* 2002;99:6292–6297.
4. Kass DJ, Kaminski N. Evolving genomic approaches to idiopathic pulmonary fibrosis: moving beyond genes. *Clin Transl Sci* 2011;4:372–379.
5. Baumgartner KB, Samet JM, Stidley CA, Colby TV, Waldron JA. Cigarette smoking: a risk factor for idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 1997;155:242–248.
6. Yang IV, Schwartz DA. Epigenetics of idiopathic pulmonary fibrosis. *Transl Res* 2014;pii: S1931-5244(14)00117-0.
7. Kim DH, Nelson HH, Wiencke JK, Zheng S, Christiani DC, Wain JC, Mark EJ, Kelsey KT. p16(INK4a) and histology-specific methylation of CpG islands by exposure to tobacco smoke in non–small cell lung cancer. *Cancer Res* 2001;61:3419–3424.
8. Wan ES, Qiu W, Baccarelli A, Carey VJ, Bacherman H, Rennard SI, Agusti A, Anderson W, Lomas DA, Demeo DL. Cigarette smoking behaviors and time since quitting are associated with differential DNA methylation across the human genome. *Hum Mol Genet* 2012;21:3073–3082.

9. Philibert RA, Sears RA, Powers LS, Nash E, Bair T, Gerke AK, Hassan I, Thomas CP, Gross TJ, Monick MM. Coordinated DNA methylation and gene expression changes in smoker alveolar macrophages: specific effects on VEGF receptor 1 expression. *J Leukoc Biol* 2012; 92:621–631.
10. Buro-Auriemma LJ, Salit J, Hackett NR, Walters MS, Strulovici-Barel Y, Staudt MR, Fuller J, Mahmoud M, Stevenson CS, Hilton H, *et al*. Cigarette smoking induces small airway epithelial epigenetic changes with corresponding modulation of gene expression. *Hum Mol Genet* 2013;22:4726–4738.
11. Tennis MA, Vanscoyk MM, Wilson LA, Kelley N, Winn RA. Methylation of Wnt7a is modulated by DNMT1 and cigarette smoke condensate in non–small cell lung cancer. *PLoS One* 2012;7:e32921.
12. Coward WR, Watts K, Feghali-Bostwick CA, Knox A, Pang L. Defective histone acetylation is responsible for the diminished expression of cyclooxygenase 2 in idiopathic pulmonary fibrosis. *Mol Cell Biol* 2009;29:4325–4339.
13. Coward WR, Watts K, Feghali-Bostwick CA, Jenkins G, Pang L. Repression of IP-10 by interactions between histone deacetylation and hypermethylation in idiopathic pulmonary fibrosis. *Mol Cell Biol* 2010;30:2874–2886.
14. Sanders YY, Tollefsbol TO, Varisco BM, Hagood JS. Epigenetic regulation of thy-1 by histone deacetylase inhibitor in rat lung fibroblasts. *Am J Respir Cell Mol Biol* 2011;45:16–23.
15. Cisneros J, Hagood J, Checa M, Ortiz-Quintero B, Negreros M, Herrera I, Ramos C, Pardo A, Selman M. Hypermethylation-mediated silencing of p14(ARF) in fibroblasts from idiopathic pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol* 2012;303:L295–L303.
16. Hu B, Gharaee-Kermani M, Wu Z, Phan SH. Epigenetic regulation of myofibroblast differentiation by DNA methylation. *Am J Pathol* 2010; 177:21–28.
17. Hu B, Gharaee-Kermani M, Wu Z, Phan SH. Essential role of MeCP2 in the regulation of myofibroblast differentiation during pulmonary fibrosis. *Am J Pathol* 2011;178:1500–1508.
18. Rabinovich EI, Kapetanaki MG, Steinfeld I, Gibson KF, Pandit KV, Yu G, Yakhini Z, Kaminski N. Global methylation patterns in idiopathic pulmonary fibrosis. *PLoS One* 2012;7:e33770.
19. Sanders YY, Ambalavanan N, Halloran B, Zhang X, Liu H, Crossman DK, Bray M, Zhang K, Thannickal VJ, Hagood JS. Altered DNA methylation profile in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2012;186:525–535.
20. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, *et al*. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* 2009;41: 178–186.
21. Ji H, Ehrlich LI, Seita J, Murakami P, Doi A, Lindau P, Lee H, Aryee MJ, Irizarry RA, Kim K, *et al*. Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature* 2010;467: 338–342.
22. Kim K, Doi A, Wen B, Ng K, Zhao R, Cahan P, Kim J, Aryee MJ, Ji H, Ehrlich LI, *et al*. Epigenetic memory in induced pluripotent stem cells. *Nature* 2010;467:285–290.
23. Ladd-Acosta C, Aryee MJ, Ordway JM, Feinberg AP. Comprehensive high-throughput arrays for relative methylation (CHARM). *Curr Protoc Hum Genet* 2010;Chapter 20:Unit.20.1.1–19.
24. Yang I, Pedersen B, Hennessy C, Davidson E, Juan Guardela B, Tedrow J, Correll M, Geraci M, Quackenbush J, Sciurba F, *et al*. Methyl-eQTL analysis in idiopathic pulmonary fibrosis (IPF) [abstract]. *Am J Respir Crit Care Med* 2013;187:A5756.
25. Demedts M, Costabel U. ATS/ERS international multidisciplinary consensus classification of the idiopathic interstitial pneumonias. *Eur Respir J* 2002;19:794–796.
26. Irizarry RA, Ladd-Acosta C, Carvalho B, Wu H, Brandenburg SA, Jeddeloh JA, Wen B, Feinberg AP. Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Res* 2008;18: 780–790.
27. Marmai C, Sutherland RE, Kim KK, Dolganov GM, Fang X, Kim SS, Jiang S, Golden JA, Hoopes CW, Matthay MA, *et al*. Alveolar epithelial cells express mesenchymal proteins in patients with idiopathic pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol* 2011;301:L71–L78.
28. Aryee MJ, Wu Z, Ladd-Acosta C, Herb B, Feinberg AP, Yegnasubramanian S, Irizarry RA. Accurate genome-scale percentage DNA methylation estimates from microarray data. *Biostatistics* 2011;12:197–210.
29. Stegle O, Parts L, Piipari M, Winn J, Durbin R. Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. *Nat Protoc* 2012;7: 500–507.
30. Stegle O, Parts L, Durbin R, Winn J. A Bayesian framework to account for complex non-genetic factors in gene expression levels greatly increases power in eQTL studies. *PLoS Comput Biol* 2010;6: e1000770.
31. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;3:Article3.
32. Pedersen BS, Schwartz DA, Yang IV, Kechris KJ. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated *P*-values. *Bioinformatics* 2012;28:2986–2988.
33. Pedersen BS, Yang IV, De S. CruzDB: software for annotation of genomic intervals with UCSC genome-browser database. *Bioinformatics* 2013;29:3003–3006.
34. 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.
35. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* 2012;28:1353–1358.
36. Fingerlin TE, Murphy E, Zhang W, Peljto AL, Brown KK, Steele MP, Loyd JE, Cosgrove GP, Lynch D, Groshong S, *et al*. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet* 2013;45:613–620.
37. Noth I, Zhang Y, Ma S-F, Flores C, Barber M, Huang Y, Broderick SM, Wade MS, Hysi P, Scuirba J, *et al*. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med* 2013;1: 309–317.
38. Phillips RJ, Burdick MD, Hong K, Lutz MA, Murray LA, Xue YY, Belperio JA, Keane MP, Strieter RM. Circulating fibrocytes traffic to the lungs in response to CXCL12 and mediate fibrosis. *J Clin Invest* 2004;114: 438–446.
39. Scotton CJ, Krupiczojc MA, Königshoff M, Mercer PF, Lee YC, Kaminski N, Morser J, Post JM, Maher TM, Nicholson AG, *et al*. Increased local expression of coagulation factor X contributes to the fibrotic response in human and murine lung injury. *J Clin Invest* 2009; 119:2550–2563.
40. Henderson WR Jr, Chi EY, Ye X, Nguyen C, Tien YT, Zhou B, Borok Z, Knight DA, Kahn M. Inhibition of Wnt/beta-catenin/CREB binding protein (CBP) signaling reverses pulmonary fibrosis. *Proc Natl Acad Sci USA* 2010;107:14309–14314.
41. Malli F, Koutsokera A, Paraskeva E, Zakynthinos E, Papagianni M, Makris D, Tsilioni I, Molyvdas PA, Gourgoulianis KI, Daniil Z. Endothelial progenitor cells in the pathogenesis of idiopathic pulmonary fibrosis: an evolving concept. *PLoS One* 2013;8:e53658.
42. Lappi-Blanco E, Lehtonen ST, Sormunen R, Merikallio HM, Soini Y, Kaarteenaho RL. Divergence of tight and adherens junction factors in alveolar epithelium in pulmonary fibrosis. *Hum Pathol* 2013;44: 895–907.
43. Hayashi H, Sano H, Seo S, Kume T. The Foxc2 transcription factor regulates angiogenesis via induction of integrin beta3 expression. *J Biol Chem* 2008;283:23791–23800.
44. Henderson NC, Sheppard D. Integrin-mediated regulation of TGFbeta in fibrosis. *Biochim Biophys Acta* 2013;1832:891–896.
45. Wolfe MS. Structure, mechanism and inhibition of gamma-secretase and presenilin-like proteases. *Biol Chem* 2010;391:839–847.
46. Misquitta-Ali CM, Cheng E, O'Hanlon D, Liu N, McGlade CJ, Tsao MS, Blencowe BJ. Global profiling and molecular characterization of alternative splicing events misregulated in lung cancer. *Mol Cell Biol* 2011;31:138–150.
47. Mousseau DD, Chapelsky S, De Crescenzo G, Kirkitadze MD, Magoon J, Inoue S, Teplow DB, O'Connor-McCourt MD. A direct interaction between transforming growth factor (TGF)-betas and amyloid-beta protein affects fibrillogenesis in a TGF-beta receptor–independent manner. *J Biol Chem* 2003;278:38715–38722.

48. Auerbach RK, Chen B, Butte AJ. Relating genes to function: identifying enriched transcription factors using the ENCODE ChIP-Seq significance tool. *Bioinformatics* 2013;29:1922–1924.
49. Wang C, Liu Z, Woo CW, Li Z, Wang L, Wei JS, Marquez VE, Bates SE, Jin Q, Khan J, *et al*. EZH2 Mediates epigenetic silencing of neuroblastoma suppressor genes CASZ1, CLU, RUNX3, and NGFR. *Cancer Res* 2012;72:315–324.
50. Seibold MA, Smith RW, Urbanek C, Groshong SD, Cosgrove GP, Brown KK, Schwarz MI, Schwartz DA, Reynolds SD. The idiopathic pulmonary fibrosis honeycomb cyst contains a mucocilary pseudostratified epithelium. *PLoS One* 2013;8:e58658.
51. Jirtle RL, Skinner MK. Environmental epigenomics and disease susceptibility. *Nat Rev Genet* 2007;8:253–262.
52. Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, Gilad Y, Pritchard JK. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol* 2011;12:R10.
53. Zhang D, Cheng L, Badner JA, Chen C, Chen Q, Luo W, Craig DW, Redman M, Gershon ES, Liu C. Genetic control of individual differences in gene-specific methylation in human brain. *Am J Hum Genet* 2010;86:411–419.
54. Yang IV, Schwartz DA. Epigenetic control of gene expression in the lung. *Am J Respir Crit Care Med* 2011;183:1295–1301.
55. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 2012;13:484–492.
56. Amin NM, Gibbs D, Conlon FL. Differential regulation of casz1 protein expression during cardiac and skeletal muscle development. *Dev Dyn* 2014;243:948–956.
57. Charpentier MS, Christine KS, Amin NM, Dorr KM, Kushner EJ, Bautch VL, Taylor JM, Conlon FL. CASZ1 promotes vascular assembly and morphogenesis through the direct regulation of an EGFL7/RhoA-mediated pathway. *Dev Cell* 2013;25:132–143.
58. Nichol D, Stuhlmann H. EGFL7: a unique angiogenic signaling factor in vascular development and disease. *Blood* 2012;119:1345–1352.
59. Di Ruscio A, Ebralidze AK, Benoukraf T, Amabile G, Goff LA, Terragni J, Figueroa ME, De Figueiredo Pontes LL, Alberich-Jorda M, Zhang P, *et al*. DNMT1-interacting RNAs block gene-specific DNA methylation. *Nature* 2013;503:371–376.
60. Yu DH, Waterland RA, Zhang P, Schady D, Chen MH, Guan Y, Gadkari M, Shen L. Targeted p16(Ink4a) epimutation causes tumorigenesis and reduces survival in mice. *J Clin Invest* 2014;124:3708–3712.
61. Kadara H, Fujimoto J, Yoo SY, Maki Y, Gower AC, Kabbout M, Garcia MM, Chow CW, Chu Z, Mendoza G, *et al*. Transcriptomic architecture of the adjacent airway field cancerization in non–small cell lung cancer. *J Natl Cancer Inst* 2014;106:dju004.