Introduction

Organ function depends upon a connective tissue stromal network that imparts topographic integrity by precisely ordering cellular and tissue compartments. The stroma consists of cells and their extracellular matrix (ECM) products. Fibroblasts maintain stromal homeostasis by integrating signals from the ECM that control their function and fate (1–4). Stromal pathology, characterized by pathological fibroblasts and a stiff ECM, is the signature of human diseases ranging from cardiovascular and pulmonary disorders to cancer (5–8), and yet we have only a limited understanding of the mechanisms governing the transition from healthy to diseased stroma (9, 10). One central unresolved issue in fibrotic disorders is the extent to which disease progression results from intrinsically diseased fibroblasts, a diseased ECM that corrupts otherwise normal fibroblasts, or a collaboration between the two.

Idiopathic pulmonary fibrosis (IPF) is a model fibrotic disease in which progressive scarring of the lungs leads to death by asphyxiation (11). It is distinguished by the aberrant proliferation of activated fibroblasts and pathological remodeling of the ECM (12). The disease process begins at the base and periphery of the lung in a subpleural distribution (13) and progresses in an upward and inward manner by contiguous spreading to form an uninterrupted fibrotic reticulum that decreases lung compliance and impairs gas exchange (14, 15). In IPF, both the fibroblasts and the ECM are profoundly altered. Lung fibroblasts obtained from IPF patients manifest aberrant activation of growth factor signaling pathways, relaxation of cell cycle controls, and alterations in apoptosis regulation (16–19). The ECM in IPF is an inelastic environment containing excessive amounts of connective tissue, matrix metalloproteinases, growth factors, and morphogens (20–23). Although it is clear that the ECM plays a crucial role in directing cellular processes and maintaining tissue integrity (24–26), the extent to which a pathological ECM can drive a pathological cellular phenotype in IPF is not well defined.

Pathological remodeling of the extracellular matrix (ECM) by fibroblasts leads to organ failure. Development of idiopathic pulmonary fibrosis (IPF) is characterized by a progressive fibrotic scarring in the lung that ultimately leads to asphyxiation; however, the cascade of events that promote IPF are not well defined. Here, we examined how the interplay between the ECM and fibroblasts affects both the transcriptome and translome by culturing primary fibroblasts generated from IPF patient lung tissue or nonfibrotic lung tissue on decellularized lung ECM from either IPF or control patients. Surprisingly, the origin of the ECM had a greater impact on gene expression than did cell origin, and differences in translational control were more prominent than alterations in transcriptional regulation. Strikingly, genes that were translationally activated by IPF-derived ECM were enriched for those encoding ECM proteins detected in IPF tissue. We determined that genes encoding IPF-associated ECM proteins are targets for miR-29, which was downregulated in fibroblasts grown on IPF-derived ECM, and baseline expression of ECM targets could be restored by overexpression of miR-29. Our data support a model in which fibroblasts are activated to pathologically remodel the ECM in IPF via a positive feedback loop between fibroblasts and aberrant ECM. Interrupting this loop may be a strategy for IPF treatment.

Conflict of interest: The authors have declared that no conflict of interest exists.

Citation for this article: J Clin Invest. 2014;124(4):1622–1635. doi:10.1172/JCI71386.
sion. By searching for shared RNA regulatory elements among the translationally activated ECM genes, we identified miR-29 as a potential regulator of the feedback loop. ECM targets of miR-29 were preferentially activated by the IFP ECM, miR-29c was downregulated in IPF and control fibroblasts by the IFP ECM, and ectopic overexpression of miR-29c rescued this translational activation of ECM genes. Thus, our data provide direct insight into the molecular underpinnings of fibrosis progression in IFP by unveiling an ECM-driven positive feedback loop that can redirect fibroblast ECM gene expression by reducing miR-29 expression, a known potent negative regulator of ECM genes.

Results

Pathological gene expression in IPF fibroblasts is primarily driven by the ECM. To explore the relative contribution of cell origin and ECM origin to gene expression patterns in fibrotic fibroblasts, we cultured IFP or control fibroblasts on IPF or control ECM in a 2 × 2 experimental design (using primary fibroblasts from 5 patients with IFP and 5 control patients) (Figure 1A). From each sample, we measured levels of both steady-state and polysome-associated RNA (Figure 1B). To assess whether this in vitro model reflects fibroblast gene expression in vivo, we compared the steady-state RNA profiles from control and IFP fibroblasts on their native ECM (IFP fibroblasts on IPF ECM compared with control fibroblasts on control ECM) with a published dataset examining ex vivo uncultured fibroblasts (29). We tested the hypothesis that those genes up- or downregulated in IFP fibroblasts on IFP ECM relative to control fibroblasts on control ECM are similarly regulated in ex vivo IPF fibroblasts by calculating Cohen’s k statistic. This analysis indicated only slight agreement between the two datasets, but greater than would be expected by chance (k = 0.053, 95% CI: 0.032–0.075). To further address whether the gene expression observed in our model system corresponds to that observed in vivo, we tested the hypothesis that genes upregulated or downregulated in IPF fibroblasts on IPF ECM are similarly regulated in a published dataset comparing RNA extracted from snap-frozen IPF tissue with RNA extracted from control tissue (30). Notably, the tissue studied contains many cell types, which could potentially dilute differences in gene expression that are specific to fibroblasts. Again, we observed marginal agreement between our data and the in vivo data (k = 0.026, 95% CI: 0.000–0.053). However, even the modest agreement between our data and two external datasets (one ex vivo and the other in vivo) — despite technical limitations due to different cell isolation procedures and a diluted fibroblast signature — supports the idea that our model dataset captures some of the disease-relevant differences in IPF fibroblast gene expression that occur in vivo.

We next focused on our data from polysome-associated RNA, because data on mRNAs that are translated more closely correspond to protein levels compared with data on steady-state mRNAs (27, 31). As expected, the analysis indicated substantial differences between IPF and control cells on their native ECM as judged by an enrichment of low P values (Figure 1C). To determine whether the observed differences were due to fibroblast-intrinsic changes or whether they resulted from the fibroblast receiving cues from a fibrotic ECM, we compared the effect of ECM type (independent of cell origin) with the effect of fibroblast origin (independent of ECM type). Strikingly, more genes were modulated by ECM type than by cell origin (compare Figure 1, D and E; P < 2.2 × 10⁻¹⁶ using the Kolmogorov-Smirnov [KS] test). After adjustment of the P values using the Benjamini-Hochberg false discovery rate (FDR) procedure (32), 389 genes showed statistically significant (FDR < 0.2) differences in the level of polysome-associated RNA between ECM types (Supplemental Table 1; supplemental material available online with this article; doi:10.1172/JCI71386DS1). At this FDR threshold, we found that only one gene was linked to cell origin. Note that this is a relatively permissive FDR threshold, allowing up to 20% false-positives. Although this is not a stringent criterion for identifying particular genes, this analysis revealed that the majority of changes we observed genome wide were driven by ECM type. We obtained similar results when the same analysis was applied to data from steady-state RNA: most of the observed differences were due to ECM type and not to cell origin (Figure 1, F–H; P < 2.2 × 10⁻¹⁶ using the KS test).

We found it intriguing that so few genes were linked to cell origin. To determine whether such genes are truly differentially expressed, we compared our results to cell origin–linked genes from one of our prior studies (conducted on synthetic type I collagen matrices) (28). We reasoned that if genes were modulated in the same direction in IPF fibroblasts across both studies, the observed differences were likely genuine. The set of genes with a P value less than 0.05 in both studies (72 genes) indicates moderate agreement between the two studies (k = 0.476, 95% CI: 0.269–0.682) (Figure 1I) and supports the supposition that there is indeed a set of genes that are differentially expressed depending on cell origin. To further investigate this issue, we analyzed a published microarray dataset that focused on steady-state RNA differences between IPF and control cells (conducted on tissue culture plastic and therefore expected to be different from the other studies conducted on biological matrices) (33). This dataset also showed only weak evidence of differences between the IPF and control cells (Figure 1J). Thus, the consensus among these studies is that although there are bona fide differences in gene expression between IPF and control fibroblasts, these differences are relatively small compared with the changes induced by the IFP ECM that we observed here.

A principal role for translation in the gene expression response to IPF ECM. Because we measured both total steady-state RNA and polysome-associated RNA, we had the opportunity to determine the relative contribution of steady-state RNA levels (reflecting transcription and RNA stability) and translation to the observed ECM-induced gene expression changes. To study translation, levels of polysome-associated RNA have to be corrected for the contribution from other mechanisms in the gene expression pathway (34). This is necessary because genes whose steady-state RNA levels increase will, by mass action alone, also show more RNA associated with polysomes without any regulation of translation. Correction is often done by considering the translational efficiency (TE) ratio (polysome-associated RNA/steady-state RNA). However, this procedure erroneously assumes that an increase in steady-state RNA is noiselessly associated 1:1 with an increase in polysome-associated RNA (leading to increased levels of false-positives and false-negatives when TE is used to compare conditions). We therefore used a more robust correction method called ANOTA (analysis of translational activity) (34, 35), which was recently found to better reflect differences in proteomes as compared with TE measures (36). In this method, polysome-associated RNA data are corrected using a linear fit to steady-state RNA data, which allows for the assessment of biological effects at the level of translation. We followed this
Figure 1
Pathological gene expression in fibrotic fibroblasts is primarily governed by diseased ECM. (A) Experimental design. IPF or control fibroblasts were cultured on IPF or control ECM using a 2 × 2 experimental design. (B) Isolation of polysome-associated RNA. Sample polysome tracing shows absorbance at 254 nm across the sucrose gradient. Transcripts that are translated are enriched among polysome-associated RNA. The portion of the gradient collected to measure polysome-associated RNA levels is indicated. (C–H) Histograms of gene-by-gene P values for different biological comparisons. Dashed lines represent theoretical null distributions. (C–E) Comparison of polysome-associated RNA levels of (C) IPF fibroblasts seeded on IPF ECM and control fibroblasts seeded on control ECM, (D) IPF fibroblasts and control fibroblasts (independent of ECM type), and (E) IPF and control ECM (independent of cell origin). (F–H) Same comparisons as in (A–C) using steady-state RNA data. (I) Genes that showed cell origin modulation in polysome-associated RNA (P < 0.05) in both our current study and in a previous study (microarray data obtained from GSE11196) were collected. Plotted are the fold changes in the previous study (x axis) and in our current dataset (y axis). Genes upregulated in one study are more likely to be upregulated in the other (κ = 0.476, 95% CI = [0.269, 0.682]). (J) Histogram of P values for comparison of RNA obtained from control and IPF cells grown under standard culture conditions (microarray data obtained from GEO GSE10921).
method and fit a per-gene linear model to determine the effect of the fibrotic ECM on translation:

$$E_{\text{poly}} = y_{\text{cell line}} \cdot E_{\text{cell}} + y_{\text{ECM}} + y_{\text{within-subject}}$$

(Equation 1)

where $E_{\text{poly}}$ and $E_{\text{cell}}$ are the data from polysome-associated and steady-state RNA, respectively, $y_{\text{cell}}$ is the fitted linear dependence of the polysome-associated RNA on the steady-state RNA, $y_{\text{ECM}}$ is the expression difference of cells on the IPF ECM compared with control ECM, $y_{\text{cell line}}$ is the relative expression of the gene in a particular cell line (included to account for the significant variation in gene expression between fibroblasts from different patients), and $e_{\text{within-subject}}$ is the within-subject error. The size (and statistical significance) of $y_{\text{ECM}}$ reflects the difference in translation induced by IPF ECM. A similar model was used to analyze the cell effects.

In order to determine whether the effects of the fibrotic ECM were primarily due to changes in steady-state RNA levels or to changes in translation, we calculated a cumulative probability distribution of the FDRs for each level (steady-state RNA, polysome-associated RNA and translation [i.e., ANOTA-corrected polysome-associated RNA]) (Figure 2A). As expected, data from polysome-associated RNA (which, in principle, combines the effects from both steady-state RNA levels and translation) showed the greatest number of significantly altered genes. Strikingly, more genes showed low FDRs when we analyzed their translation than did steady-state RNA levels ($P = 2.6 \times 10^{-8}$ using the KS test). As expected from the low number of genes whose expression depended on cell origin (Figure 1, D and G), the same analysis of the cell origin variable identified very few genes, even at relaxed FDR thresholds (Figure 2B). Thus, we found that the majority of genes that are affected by residence on the diseased ECM were modulated at the level of translation.

As a technical validation step, we next determined whether the observed predominance of translational control among the ECM-induced genes was due to larger fold change differences in data from polysome-associated RNA and not to smaller variances (the latter would indicate that larger measurement noise masked differential expression in steady-state RNA data). A comparison of the fold changes corroborated that the ECM-induced differences in the polysome-associated data are typically higher than in the steady-state RNA data (Figure 2C). The associated variances (residual sum of squares) were also typically higher (Figure 2D). This is expected, because the preparation of polysome-associated RNA involves additional processing, which increases measurement noise. That we found more significant genes despite higher variance is strong evidence that the dominance of translational control is genuine.

Although the majority of changes observed in response to IPF ECM were at the level of translation, there were also genes whose transcript level changed. To assess whether regulation converged on the same genes, we compiled all genes that showed modulation either through translation or at the steady-state RNA level and looked for overlap (FDR < 0.3, the lowest threshold at which a subset of genes was altered at the steady-state RNA level). A comparison of these genes generated four distinct modules: those genes up- or downregulated either by translation or by steady-state RNA, with only one gene, PRIC285, overlapping (Figure 2E). Despite the low number of genes modulated by steady-state RNA levels, this suggests that translation and transcription/RNA stability act independently to induce changes to the gene expression profile under these conditions.

Fibrotic ECM activates translation of ECM genes. When we used a gene set enrichment approach (37) to determine whether groups of genes that are associated with a given biological function were translationally upregulated by the IPF ECM, we surprisingly identified several ECM ontologies (e.g., “ECM region,” “cell adhesion,” and “collagen,” see Table 1). This appeared to be of significance because of the profound changes that the ECM undergoes during IPF disease progression. We therefore further characterized the expression of ECM genes by examining the regulation of all genes in the “ECM region” ontology using data for steady-state RNA or translation. As expected, such genes showed coordinated translational activation by the IPF ECM (Figure 3A, upper panel) and were under stronger regulation than non-ECM genes (Figure 3A, lower panel). Similarly, but to a lesser extent, genes in the ECM region ontology were more translationally regulated by cell origin than non-ECM genes (Figure 3B). The same analysis using data for steady-state RNA indicated that ECM region genes were not regulated more than non-ECM genes by ECM origin (Figure 3C) and that although ECM region genes were more likely to be regulated by cell origin at the steady-state RNA level than non-ECM genes, this regulation was not directional: genes were both up- and downregulated (Figure 3D).

This raises the possibility that ECM origin and cell origin together determine the expression pattern for ECM genes by regulating overlapping gene subsets either through changes in translation or in steady-state RNA levels. To assess this, we collected all ECM region genes that were significantly altered (P < 0.05) in at least one of the four comparisons (i.e., regulated by cell origin or ECM origin either at the level of steady-state RNA or translation). To separate genes based on which comparisons showed modulation, we used k-means clustering of signed (i.e., upregulation [+ ] or downregulation [−]) -log_{10} P values (Figure 4A). As expected from our finding that changes in translational activity and steady-state RNA levels target different subsets of genes (Figure 2E), this clustering revealed a clear segregation of genes into a “translation profile” and a “steady-state RNA profile,” with only a small number of genes that did not fit cleanly into either profile. Notably, at the level of translation, most genes that were regulated by cell origin were also regulated by ECM origin, indicating a convergence of cell- and ECM-origin effects for a subset of genes at this control level. In contrast, at the level of steady-state RNA, the majority of genes regulated by cell origin were not regulated by ECM origin. Because the clustering indicated that genes are regulated by either cell origin, ECM origin, or both cell and ECM origin, we compared the fold changes in each of these subgroups. As expected, in the translation profile the cell-regulated genes (Figure 4B, left) and the ECM-regulated genes (Figure 4B, center) had larger cell effects and ECM effects, respectively. In the module of coregulated genes, the ECM and cell effects were similar (Figure 4B, right), indicating that both ECM origin and cell origin shape the expression pattern of these genes to a comparable extent. Of note, gene ontologies associated with ECM remodeling (i.e., “ECM disassembly” and “regulation of angiogenesis”) were found to be overrepresented in the ECM-regulated genes, while the two other subgroups had too few genes to allow for efficient analysis of enrichment. The same analysis of fold changes for genes regulated at the steady-
state RNA level mirrored our findings for the translationally regulated group (Figure 4C) and notably identified an “inflammatory response” ontology dependent on ECM origin.

Thus, at each regulatory level (i.e., steady-state RNA and translation), we observed ECM genes that were regulated independently by the ECM and by cell origin. In addition, we observed convergence on a common gene set at the translation level. This implies that although the primary effect in our system is activation of translation induced by the IPF ECM, cell-intrinsic mechanisms also influence the expression of ECM region genes.

ECM genes that are translationally activated by IPF ECM comprise the IPF ECM in vivo. These in vitro results suggest that fibroblasts coordinately induce translation of a set of ECM genes primarily in response to the diseased ECM. It was therefore important to determine whether these ECM genes are found in the IPF ECM in vivo. Our recent proteomic analysis of decellularized IPF ECM

---

**Figure 2**
Diseased ECM predominantly affects gene expression by modulating translation. (A and B) Cumulative FDR probability distributions for three control levels: polysome-associated RNA, steady-state RNA, and translation (ANOTA-corrected polysome-associated RNA). Each line indicates the fraction of genes (y axis) that passed a given FDR threshold (x axis) for each control level. (A) Comparison between IPF and control ECM. (B) Comparison between IPF and control fibroblasts. (C) Density plot of gene-by-gene fold changes from the ECM origin comparison. (D) Density plot of gene-by-gene variances in comparisons of ECM origin. (E) Heatmap of genes that were upregulated (yellow) or downregulated (blue) at the steady-state RNA or translation level (FDR < 0.3). P<, probability. Abs, absolute value; MSS, mean sum of squares.
identified a panel of proteins that constitute the ECM in IPF. We collected all such genes (termed here IPF-detected proteins) and determined whether they were selectively regulated by the IPF ECM. Indeed, there appeared to be a broad translational upregulation of genes detected in vivo in the IPF ECM in response to decellularized IPF ECM (Figure 5A). To quantify this, we assessed cumulative P value distributions of different nonoverlapping gene subsets (Figure 5, B and C). At the level of translation, we found that ECM genes were more translationally activated by the IPF ECM than all other genes (seen in Figure 3A as well). Strikingly, IPF-detected proteins were in turn significantly more translated than other ECM genes (P = 3.9 × 10⁻⁵, 1-tailed Mann-Whitney U test). A similar analysis for the cell origin comparison showed that IPF-detected proteins were not more regulated by cell origin than other ECM genes. These findings strongly indicate that the IPF ECM stimulates increased translation of proteins, which are themselves part of the IPF ECM. Moreover, in agreement with identification of ECM origin as the major determinant for disease gene expression, although cell origin has an effect on the synthesis of ECM proteins, these proteins do not appear to be enriched in the IPF ECM in vivo.

A positive feedback loop between the IPF ECM and fibroblasts involves miR-29c. Our findings are in agreement with the idea that there is a positive feedback loop, wherein the pathological ECM translationally activates ECM genes. There are many mechanisms that can regulate translation (39), including microRNAs, which affect translation and/or RNA stability and thereby protein levels. To search for the possible molecular mechanisms underlying this positive feedback loop, we determined whether the predicted mRNA targets of any microRNA family were overrepresented among the translationally activated ECM genes (those genes with a P value less than 0.05). Indeed, there was one microRNA family (miR-29abcd) whose targets were significantly overrepresented in the group of ECM genes that were translationally activated by the IPF ECM (P = 6.28 × 10⁻⁵, Fisher’s exact test). This aligns with recent reports associating miR-29 with IPF (40, 41). Indeed, targets of miR-29 were abundant among the IPF-detected proteins that were translationally activated by the IPF ECM (Figure 5A). Our detailed analysis showed that among all IPF-detected proteins, those containing miR-29 targets were significantly more likely to be translationally regulated by the IPF ECM than those lacking targets (P = 0.02, 1-tailed Mann-Whitney U test; Figure 5B). In contrast, we did not observe any cell origin regulation by miR-29 (Figure 5C). Moreover, we did not find regulation of IPF-detected proteins or miR-29 targets at the steady-state RNA level (Supplemental Figure 1), suggesting that translation is the primary mechanism for regulation downstream of miR-29 under these conditions.

These data indicate that reduced expression of miR-29 in response to IPF ECM could mechanistically explain the positive feedback loop between fibroblasts and the pathological ECM (because microRNAs negatively regulate gene expression). To investigate this, we quantified the levels of the three human forms of the miR-29 family (miR-29a, miR-29b, and miR-29c) in vitro using the 2 × 2 experimental model (Figure 1A). Similarly to our analysis in Figure 5, B and C, we compared both the effect of IPF ECM (independent of cell origin) and IPF cell origin (independent...
Consistent with the increased translation observed for IPF-detected proteins on the IPF ECM, we found that the level of miR-29c was reduced ($P = 0.031$) and that the other two miR-29 forms trended downward on the IPF ECM (Figure 5D). Moreover, consistent with the lack of any regulation of miR-29 by cell origin, none of the three miR-29 forms were significantly regulated between the two cell origins (Figure 5E). Thus, taken together with our genome-wide data indicating that miR-29 targets are upregulated, modulation of miR-29 expression may partly explain the increased translation of ECM genes on the IPF ECM.

Figure 4
Converging and independent modulation of ECM gene translation is dependent upon ECM and cell origins. (A) Translation (ANOTA-corrected) and steady-state RNA profiles of genes in the ECM region gene ontology that were differentially expressed ($P < 0.05$) in any comparison. Values are log$_{10}$ $P$ values (yellow denotes upregulation in IPF, blue denotes downregulation). (B) Close-up of the translation profile. Genes are divided into three categories: cell-regulated, ECM-regulated, and coregulated. Density plots of the absolute fold changes induced by each biological variable are shown. Selected gene ontologies that are overrepresented ($P < 0.01$, calculated using Fisher's exact test) are shown (see Supplemental Table 2 for the complete list). There were no significantly overrepresented gene ontologies in the cell-regulated group. (C) Same analysis for the steady-state RNA profile.
Figure 5
A positive feedback loop between diseased ECM and the fibroblast involves modulation of miR-29 expression. (A) Translation profile of genes whose protein products are present in IPF lung ECM. Shown are –log_{10} P values from the cell origin and ECM origin comparison (yellow denotes upregulation in IPF, blue denotes downregulation). miR-29 targets are designated by solid bars above the heatmap. (B and C) Cumulative P value probability distributions for (B) ECM origin and (C) cell origin are shown. Dotted line represents the theoretical null distribution. Black line represents all genes except ECM region genes. Red line represents all ECM genes except those detected in the IPF lung (KS P value for comparison with “All genes”). Blue line represents all genes from the IPF lung except miRNA-29 targets (KS P value for comparison with “ECM genes”). Yellow line represents all IPF-detected miR-29 targets (KS P value for comparison with “IPF-detected proteins”). (D and E) Levels of miR-29 species were quantified using qPCR. In all comparisons, the arbitrary units were normalized to the control level. The levels of miR-29c between control and IPF ECM were significantly altered (P = 0.031). It should be noted that the ECM comparison is paired so the single error bar represents the standard error between the paired differences.
or that miR-29 may be responsible for some of the upregulation. This indicated incompleteness of the miR-29 annotation and/or that not being annotated as harboring any miR-29 targets (Figure 6E). Returning toward baseline when miR-29c was overexpressed despite genes upregulated by the IPF ECM did not return to baseline expressed when cultured on IPF ECM (Figure 6B), whereas cells treated with control virus showed aberration of miR-29 could rescue the translation of ECM genes in IPF fibroblasts. For this, IPF fibroblasts infected with an lentivirus expressing either an hsa-miR29c construct or a scrambled construct on IPF ECM (Figure 5D). Whereas cells treated with control virus showed aberration of miR-29 is known to activate the synthesis of ECM products (47) the IPF ECM feedback loop; miR-29c overexpression can abrogate pathological upregulation of miR-29–targeted ECM genes induced by IPF ECM.

**Discussion**

Despite the profound changes that the ECM undergoes in fibrosis, there has been a paucity of studies exploring how these changes affect fibroblast gene expression genome wide. Here, we addressed this gap by seeding fibroblasts originating from IPF or control lungs into IPF ECM or control ECM and measuring the resultant gene expression. We thereby determined changes in gene expression that were intrinsic to the IPF fibroblast and changes that resulted from the fibroblast responding to the diseased ECM. Surprisingly, ECM origin had a greater impact on gene expression than did cell origin, and the majority of these changes occurred at the translation level and not at the transcription/RNA stability level. This is in agreement with emerging data positioning translation as the principal posttranscriptional mechanism (27, 42).

**The issue of cell-intrinsic versus microenvironment-dependent gene expression is complex.** We endeavored in this study to examine genome wide how the interaction between fibroblasts and the ECM differs when the same fibroblasts are cultured on control ECM compared with when they are cultured on fibrotic ECM. The isolation of fibroblasts from the lung and their subsequent expansion, although necessary to conduct such experiments, could potentially mask differences between IPF and control fibroblasts that were present in vivo, an effect that may be partially responsible for the relatively few cell-intrinsic differences that we observed. Nonetheless, agreement with previous ex vivo and in vivo datasets indicates that the differences we did observe are likely to be trustworthy. Because of these issues, we focused on the relatively larger differences in gene expression induced by fibrotic ECM.

We found that ECM genes targeted by miR-29 were selectively upregulated in response to the fibrotic ECM, spotlighting miR-29 as a critical lynchpin in the pathological interaction between the ECM and the fibroblast. This finding is consistent with a general role for miR-29 in fibrosis; expression of this microRNA is not only low in IPF (43) but in other fibrotic diseases as well (44–46). In fact, a recent study has shown that miR-29 gene transfer into the lungs of mice with bleomycin-induced lung fibrosis can arrest — but not reverse — the progression of fibrosis (47). In addition, downregulation of miR-29 is known to activate the synthesis of ECM products by lung fibroblasts (48). Together with our observation that miR-29c is downregulated by fibrotic ECM, these findings support a causal link between fibrotic ECM, downregulation of fibroblast miR-29, and the progression of fibrosis. Further study of how the cell-matrix interaction induces this miR-29 suppression could ultimately lead to potential targets for IPF treatment.

**Supporting the importance of the cell-ECM interaction in IPF,** the IPF ECM translationally activated the “integrin-binding” ontology (Table 1 and Supplemental Figure 2A). Moreover, when compiling all integrin-binding reactions from the Reactome database, we focused on the relatively larger differences in gene expression induced by fibrotic ECM.

**Table 1**

<table>
<thead>
<tr>
<th>GO term</th>
<th>FDR</th>
<th>Set size</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0005576 Extracellular region</td>
<td>1.02 × 10^{-12}</td>
<td>491</td>
</tr>
<tr>
<td>GO:0005615 Extracellular space</td>
<td>3.47 × 10^{-4}</td>
<td>238</td>
</tr>
<tr>
<td>GO:0031012 ECM</td>
<td>2.38 × 10^{-3}</td>
<td>83</td>
</tr>
<tr>
<td>GO:005789 Endoplasmic reticulum membrane</td>
<td>2.13 × 10^{-4}</td>
<td>371</td>
</tr>
<tr>
<td>GO:005887 Integral to plasma membrane</td>
<td>4.21 × 10^{-4}</td>
<td>288</td>
</tr>
<tr>
<td>GO:005578 Proteinaceous ECM</td>
<td>4.37 × 10^{-4}</td>
<td>58</td>
</tr>
<tr>
<td>GO:004872 Receptor activity</td>
<td>6.45 × 10^{-4}</td>
<td>278</td>
</tr>
<tr>
<td>GO:007155 Cell adhesion</td>
<td>1.01 × 10^{-3}</td>
<td>184</td>
</tr>
<tr>
<td>GO:005178 Integrin binding</td>
<td>2.17 × 10^{-3}</td>
<td>36</td>
</tr>
<tr>
<td>GO:005201 ECM structural constituent</td>
<td>2.20 × 10^{-3}</td>
<td>27</td>
</tr>
<tr>
<td>GO:005788 Endoplasmic reticulum lumen</td>
<td>2.30 × 10^{-3}</td>
<td>60</td>
</tr>
<tr>
<td>GO:005604 Basement membrane</td>
<td>3.80 × 10^{-3}</td>
<td>33</td>
</tr>
<tr>
<td>GO:005509 Calcium ion binding</td>
<td>4.46 × 10^{-3}</td>
<td>211</td>
</tr>
<tr>
<td>GO:007040 Lysosome organization</td>
<td>1.34 × 10^{-2}</td>
<td>20</td>
</tr>
<tr>
<td>GO:005581 Collagen</td>
<td>1.58 × 10^{-2}</td>
<td>28</td>
</tr>
<tr>
<td>GO:007411 Axon guidance</td>
<td>1.87 × 10^{-2}</td>
<td>147</td>
</tr>
<tr>
<td>GO:000139 Golgi membrane</td>
<td>1.90 × 10^{-2}</td>
<td>250</td>
</tr>
<tr>
<td>GO:000665 Amino acid transport</td>
<td>1.90 × 10^{-2}</td>
<td>13</td>
</tr>
<tr>
<td>GO:0008201 Heparin binding</td>
<td>2.27 × 10^{-2}</td>
<td>46</td>
</tr>
<tr>
<td>GO:009986 Cell surface</td>
<td>2.27 × 10^{-2}</td>
<td>127</td>
</tr>
<tr>
<td>GO:005012 Receptor binding</td>
<td>2.58 × 10^{-2}</td>
<td>100</td>
</tr>
<tr>
<td>GO:005624 Membrane fraction</td>
<td>3.39 × 10^{-2}</td>
<td>226</td>
</tr>
<tr>
<td>GO:0043867 Posttranslational protein modification</td>
<td>4.38 × 10^{-2}</td>
<td>111</td>
</tr>
<tr>
<td>GO:0059090 Leukocyte migration</td>
<td>4.75 × 10^{-2}</td>
<td>47</td>
</tr>
</tbody>
</table>

The list of genes ranked by change induced in translation by ECM type (signed –log10 P value) was analyzed using GAGE (37) to identify gene ontologies that are activated through translation. GO, gene ontology.
This pattern of ECM-associated gene expression in our data raises an interesting question. Why are fibroblasts that are seeded into a collagen-rich ECM upregulating collagen, other ECM macromolecules, and their cognate integrin receptors? No previous study has shown IPF ECM to induce such an effect, nor has any study shown the antifibrotic miR-29 family to be downregulated by fibrotic ECM. This modulation provides evidence of a positive feedback loop.
involving miR-29c whereby IPF ECM induces otherwise normal fibroblasts to produce additional ECM components (Figure 7). If this interpretation is correct, it suggests a model of IPF disease progression in which some initial insult creates a small fibrotic region in the lung. This fibrotic ECM corrupts nearby fibroblasts, which further remodel the surrounding lung tissue, spreading the fibrosis. Such a model would explain the contiguous nature of the fibrotic reticulum. It should be noted, however, that our data in no way preclude the possibility that subtle intrinsic changes to the fibroblast itself could also be important to disease initiation and progression. Indeed, we found that cell origin and ECM origin cooperatively regulated translation of a common set of ECM genes. We speculate that genetic or epigenetic changes to the fibroblast after what might be a physiological signal to heal is responsible for the initial fibrotic event and also prevents the cell from interrupting this feedback loop during disease progression. It remains to be determined whether targeting this feedback loop will be a way forward for the treatment of IPF.

Methods

Primary human lung fibroblast cell lines. Human primary fibroblasts from ten different donors were used. These consisted of five control samples (mean age, 69 years) (histologically uninvolved lung distant from the resected tumor) and five samples from patients with IPF (mean age 63 years) (histologically confirmed usual interstitial pneumonitis). Tissue was obtained at the time of biopsy, autopsy, lung resection, or lung transplantation following procedures approved by the University of Minnesota Institutional Review Board for Human Subjects Research. The regions of lung chosen for IPF cells and ECM were documented by the pathologist of record to meet all criteria for usual interstitial pneumonitis, including active fibroblast foci and did not represent areas that could be described as end-stage tumor. Cells between subcultivation numbers 4 and 7 were used in this study.

Decellularized primary human lung ECM. Pieces of lung tissue measuring approximately 30 cm³ were frozen at ~70°C and attached to 24 multiwell clusters filled with ice. The plate-adhered frozen tissue was allowed to equilibrate to a temperature of ~15°C. A block plane was used to carefully make 200-μm slices of frozen tissue, which were immediately placed in tissue culture dishes containing PBS and frozen in bulk at ~20°C for future use. When thawed, tissue slices were decellularized over the course of 4 days as follows: tissue slices were placed in a 100-mm Petri dish, and 15 ml of 1% SDS in H₂O (lysis solution) was added. Petri dishes were shaken at approximately 30 oscillations per minute on an orbital shaker at room temperature for 1 hour. Lysis solution was aspirated and replaced with 15 ml of fresh lysis solution and placed again on the orbital shaker at the same settings. After an additional hour, lysis solution was replaced, and the tissue was shaken overnight at room temperature. Lysis solution was removed and replaced with 1% Triton X-100 in H₂O followed by three 1-hour cycles of agitation on an orbital shaker in replacement Triton X-100 solution. After treating with Triton X-100 overnight on an orbital shaker, tissue slices were rinsed twice with PBS and rinsed 8 times with deionized H₂O. To lyse any residual nuclei, 15 ml of 1 M NaCl was added, and the tissue was placed on an orbital shaker at room temperature for 1 hour before being rinsed once with PBS and twice with deionized water. The tissue slices were treated with DNAase (20 μg/ml) plus 4.2 mM MgCl₂ at 37°C for 1 hour. After aspirating the DNAase, the tissue was rinsed twice with deionized water and once with PBS plus antibiotics for 30 minutes at room temperature on an orbital shaker. The PBS was removed and replaced with 20 ml of complete medium (DMEM, 10% FBS, and 300 μ/l penicillin, 300 μ/l streptomycin, and 75 μ/l amphotericin B) and shaken overnight at room temperature. This decellularization procedure is similar to a previously described procedure (50). Tissue slices were incubated (37°C, 5% CO₂) for 1 hour prior to seeding with cells. Cultured fibroblasts were released from tissue culture dishes with 0.25% trypsin/EDTA, centrifuged for 5 minutes at 900 g, and resuspended in complete medium at a concentration of 10⁶ cells/ml. Each tissue slice was placed in a 50-ml conical tube with IPF or control fibroblasts (10⁵ cells/cm² of decellularized ECM). The conical tubes containing tissue slices and cells were gently rotated at 1 oscillation per minute in a humidified 37°C/10% CO₂ incubator for 18 hours with no media changes. In order to minimize the effect of variance induced by running samples at different times (batch variance), each control cell line was processed in tandem with an IPF cell line.

Following incubation, cycloheximide was added to the media at a final concentration of 100 μg/ml to immobilize ribosomes on RNA. Tissues were returned to the incubator for 5 minutes. The tissue slices were rinsed thoroughly in PBS containing cycloheximide to remove nonadherent cells. The slices were blotted semi-dry, flash-frozen in liquid nitrogen, and stored at ~70°C.
RNA isolation and polysome preparations. A mortar and pestle set was placed into liquid nitrogen to equilibrate the temperature. Flash-frozen tissue was removed from the freezer and placed into the mortar. Tissue was ground to a fine powder and scraped into a liquid nitrogen-cooled microcentrifuge tube. A portion of the powder approximating 5% of the total material was placed into a microcentrifuge tube containing 500 μl of Tri-reagent (Sigma-Aldrich) and 100 μl of chloroform. RNA was isolated and precipitated using isopropanol according to the manufacturer’s instructions.

The remainder of the frozen powder was added to an ice-cooled 7-ml Dounce homogenizer. Swelling buffer (375 μl) was immediately added (10 mM NaCl, 20 mM Tris, pH 7.5, 3 mM MgCl₂) and supplemented with 200 U of Superase-In RNase Inhibitor (Thermo Fisher Scientific) and DTT (1 mM final concentration). Lysis buffer (125 ml; 0.2 M sucrose and 1.2% Triton X-100 in low-salt buffer [LSB]) was added, and the homogenate was churned 15 times with a tight-fitting pestle. The homogenate was transferred to a microcentrifuge tube and centrifuged at 16,000 g for 1 minute. The clarified supernatant was transferred to a clean tube containing 100 U of Superase-In and centrifuged at 16,000 g for an additional 3 minutes. The supernatant was transferred to a tube containing 15 μl of 5 M NaCl and 60 U of Superase-In RNase and layered onto a 5-ml, 0.5–1.5 M sucrose gradient, which was centrifuged at 200,000 g in a Beckman Coulter SW55Ti rotor for 80 minutes at 4°C. The RNA was fractionated and prepared as previously described (51).

Protein measurement of IPF lung tissue has been previously described (38). Microarray analysis. Total and heavy RNA samples were submitted to the University of Minnesota Biomedical Genomics Center for cDNA labeling and microarray analysis using the Illumina Expression BeadChip system. Bead summary data were normalized using the robust spline normalization and log₂ transformed in R/Bioconductor. Probes for which the Bead summary data were normalized using the robust spline normalization and log₂ transformed in R/Bioconductor. Probes for which the expression level exceeded all negative controls (detected P value = 0.0) in 2 of 3 of steady-state samples were included in downstream analysis. Gene expression levels were mapped from probe levels by selecting the highest expressing probe from each mapped gene.

The experiment is a split-plot design, with cell origin and processing batch as between-subject factors and ECM type as a within-subject factor. Therefore, to analyze the cell-based effects, the expression values for both ECM types (control and IPF) were averaged across both ECM types for each cell line for each gene. To probe for expression differences in the levels of polysome and total steady-state RNA, the following per-gene linear model was fit:

(E_{IPF} + E_{control}) / 2 = γ_{batch} + γ_{cell-origin} + ε_{within-subjects} 

(Equation 2)

where $E_{IPF}$ and $E_{control}$ are the expression of a particular cell line on IPF and control ECM, respectively (averaged to probe the cell origin differences), $γ_{batch}$ accounts for variance among sample batches, $γ_{cell-origin}$ is the expression difference between IPF and control fibroblasts, and $ε_{within-subjects}$ is the between-subjects error.

ECM effects were then analyzed in a paired design; each cell line was cultured on both ECM types. Note, since for each cell line both ECM types were processed simultaneously, the processing batch is not included in the model.

Gain-of-miR-29c-function. A lentivirus plasmid expressing hsa-miR29c (PMIRH29c-PA1) was obtained from System Biosciences. HEK293T/c17 cells were purchased from ATCC and cultured in DMEM with 10% FBS and antibiotics (penicillin 100 IU/ml, streptomycin 100 IU/ml, and amphotericin B 25 μg/ml). Third-generation lentiviral packaging plasmids were extended to include a linear fit to the total steady-state RNA levels. For cell-type effects, this yielded:

$$E_{polysome} = γ_e + γ_{cell-origin} + γ_{cell-line} + ε_{within-subjects}$$

(Equation 4)

where $γ_e$ is the fitted linear dependence of the total steady-state RNA expression level on the polysome-associated expression level. Similarly, for analysis of the ECM effects, the model fitted was:

$$E_{polysome} = γ_e + γ_{IPF} + γ_{control} + γ_{cell-origin} + ε_{within-subjects}$$

(Equation 5)

where again a paired design was used. All statistics were calculated using the “car” R package (52). All P values were adjusted using random variance model (RVM) statistics (53). Gene set enrichment analysis was performed using the GAGE R package (37) on a preranked list of genes (according to their signed -log₁₀ P values) and ontologies from the Gene Ontology Consortium (54).

Genes targeted by mir-29 were those predicted by TargetScanHuman by the presence of a 7-mer or 8-mer site matching the target region of the mir-29abc6d family (55). The microarray data from this publication have been submitted to the Gene Expression Omnibus (GEO) database (GEO GSE45686).

External datasets. External datasets (GEO GSE17978, GSE10921, and GSE10667) were acquired through the GEO. Statistics were calculated from the obtained data using an equal variance Student’s t test. The obtained P values were corrected using RVM.

To check for coherence among external datasets (29, 30) and our current data, in each comparison genes were divided into four categories: those genes upregulated in IPF in both the external dataset and our data, genes downregulated in both datasets, genes upregulated in the external data and downregulated here, and genes downregulated here but upregulated in the external dataset. The resulting contingency tables were tested for agreement using Cohen’s k coefficient.

miR quantification. qPCR was performed on total RNA samples for each of the cell/ECM conditions using the miScript SYBR Green PCR kit (QIAGEN) according to the manufacturer’s instructions. qPCR was performed for miR-29a, miR-29b, and miR-29c using prevalidated primers from the kits. Amplification was allowed to proceed for 40 cycles at which time product was quantified at the log-linear portion of the curve using LightCycler analysis software.
pMDLg/RRE and pRSV-REV were obtained from Addgene. For the generation of VSV-G pseudotyped lentivirus pMRH29c (9.3 μg) along with pMD2.G (2.9 μg), pMDLg/RRE (5.4 μg) and pRSV-REV (2.3 μg) were cotransfected into 10-cm dishes of subconfluent HEK293T/c17 cells using Fugene 6 (60 μl) (56–61). DNA complexes were incubated with the cells for 4 hours (37°C, 5% CO2). Medium was replaced and replaced with 5 ml DMEM, 10% FBS, and antibiotic medium containing 2 mM caffeine (62). The transfected cells were incubated for 24 hours (37°C, 5% CO2), and media were replaced and replaced with 5 ml DMEM, 10% FBS, and antibiotic medium containing 2 mM caffeine. Viral supernatants were harvested 48 hours after transfection, filtered through a 0.45-μm-pore-size PVDF filter, aliquoted, and frozen at −80°C. Target cells were plated in 6-well dishes and allowed to attach overnight. Virus was added to cells in DMEM plus 10% FBS with polybrene at a final concentration of 8 μg/ml, and the plate was centrifuged (1,200 g, 25°C) for 1 hour and incubated for 16 hours (37°C, 5% CO2). Virus was removed, 2 ml fresh medium per well was added, and incubation was continued until the cells were used for experiments.

qPCR. Low-passage IPF cells from a single patient were grown in 6-well cluster plates to 80% confluence. Media was aspirated and replaced with either miR29c lentivirus medium or control lentivirus medium diluted 1:4 in growth medium supplemented with polybrene (8 μg/ml). The dishes were centrifuged for 60 minutes at 1,000 g, 25°C) for 1 hour and incubated for 16 hours (37°C, 5% CO2). Virus was removed, 2 ml fresh medium per well was added, and incubation was continued until the cells were used for experiments.

qPCR. The transfected cells were incubated for 24 hours (37°C, 5% CO2), and media were replaced and replaced with 5 ml DMEM, 10% FBS, and antibiotic medium containing 2 mM caffeine (62). The transfected cells were incubated for 24 hours (37°C, 5% CO2), and media were replaced and replaced with 5 ml DMEM, 10% FBS, and antibiotic medium containing 2 mM caffeine. Viral supernatants were harvested 48 hours after transfection, filtered through a 0.45-μm-pore-size PVDF filter, aliquoted, and frozen at −80°C. Target cells were plated in 6-well dishes and allowed to attach overnight. Virus was added to cells in DMEM plus 10% FBS with polybrene at a final concentration of 8 μg/ml, and the plate was centrifuged (1,200 g, 25°C) for 1 hour and incubated for 16 hours (37°C, 5% CO2). Virus was removed, 2 ml fresh medium per well was added, and incubation was continued until the cells were used for experiments.

Acknowledgments
This work was supported by NIH grants P01 HL11775 (to C.A. Henke), R01 HL089249 (to P.B. Bitterman), and F32 HL116105 (to M.W. Parker); by funds provided by the O’Brien family and the Carlson Family Fund (to M.W. Parker); and by the Swedish Research Council and the Swedish Cancer Foundation (to O. Larson). E.S. White was supported by NIH grant U01 HL111016.

Received for publication June 3, 2013, and accepted in revised form December 27, 2013.

Address correspondence to: Matthew W. Parker, 420 Delaware St. SE, Pulmonary MMC: 27C, Minneapolis, Minnesota 55455, USA. Phone: 612.626.1124; Fax: 612.625.2174; E-mail: park0523@umn.edu.